INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONES: 224281/2/3/4/5/6
CABLE: CENTRATOM - TELEX 460392-I

SMR/111 - 29

SECOND SUMMER COLLEGE IN BIOPHYSICS

30 July - 7 September 1984

Lecture 4: Hydration of DNA: position and orientation of water molecules and counter-ions (by Monte Carlo simulation).

Lecture 5: Ion selectivity and trans-membrane channels.

G. CORONGIU
IBM Corporation
IS/TG
Poughkeepsie, N.Y. 12602
U.S.A.

Lecture 1

Simulations of the Solvent Structure for Macromolecules. I. Solvation of B-DNA Double Helix at $T = 300$ K

GIORGINA CORONGIU* and ENRICO CLEMENTI, *IBM–DPPG, P.O. Box 390, Poughkeepsie, New York 12602*

Synopsis

Monte Carlo simulations are reported for a system of 447 water molecules enclosing a B-DNA double-helix fragment with 12 base pairs and the corresponding sugar and phosphate units. From a detailed analysis on the interaction energies and probability distributions (at a simulated temperature of 300 K), the water molecules can be partitioned into clusters strongly interacting with (1) the phosphates, (2) the sugars, (3) the sugars *and* the bases, and (4) the base pairs. In addition, *transgroove* and *interphosphate* filament of hydrogen-bonded water molecules have been detected. From simulations performed with variable numbers of water molecules, a theoretical isotherm has been obtained, with the characteristic sigmoidal shape, known from absorption–desorption experiments on related systems. The expected main features for the structure of water molecules solvating B-DNA with Na^+ counterions are briefly discussed at the end of the paper.

INTRODUCTION

We have previously reported on the interaction between one water molecule and the bases[1] and base pairs[2] of nucleic acids: A-DNA single helix,[3] B-DNA single[4] and double helix.[4,5] In addition, Monte Carlo simulations (at 300 K) have been presented for a cluster of water molecules enclosing the bases and the base pairs[2] or a limited region around A-DNA single helix[3] and B-DNA single helix.[4,5] These studies represent preliminary steps. We extend our previous effort by considering, via simulations, not only a much larger number of water molecules than previously, but also the effect of counterions, initially the Na^+ ion; we report on *qualitative* features of Na^+–B-DNA at 300 K double helix and on *quantitative* aspects of B-DNA in solution at 300 K. The B-DNA double helix fragment we consider has been discussed previously (see Ref. 2), and consists of 12 base pairs (namely, two more base pairs than needed to reproduce a full B-DNA double helix turn) with the corresponding sugar and PO_4^-—CH_2 units. The B-DNA double-helix fragment is enclosed in a cylinder with its axis coaxial to the B-DNA long axis (z-axis). The cylinder height is 36.0 Å with a base radius of 14.5 Å. The two base pairs and the corresponding sugar units kept outside the cylinder (one above and one below) have been added in order

* Permanent address: Istituto Ricerche G. Donegani S.p.A., Via G. Fauser 4, Novara, 28100, Italy.

Biopolymers, Vol. 20, 551–571 (1981)

CCC 0006-3525/81/030551-21$02.10

to improve the interaction field descriptions at the bottom and top ends of the B-DNA fragment. In the Monte Carlo simulation reported below, 447 water molecules have been placed into the cylinder. The equilibration process was carried out for 2×10^6 conformations; the statistical data analyzed are obtained from additional 2×10^6 Monte Carlo "moves" (these computations were carried out on an IBM 370/3033 computer).

GRAPHICAL REPRESENTATION OF THE PROBABILITY DENSITY

Some of the analysis reported here was not carried out with the use of probability density maps, as done in the past,[2-9] since these are somewhat difficult to read and even more difficult to use as input data. The probability maps have been replaced by a model based on a new algorithm described by the following four steps:

1. After computation of the probability density maps, the probability density maxima for the oxygen atoms are located (neglecting low probability maxima by selecting a threshold that ensures limiting the number of maxima to 447, the number of water molecules enclosed into the cylinder).
2. For the hydrogen atoms, the probability maxima are determined subject to the constraint of being located on a sphere of radius equal to the O—H internuclear separation in H_2O.
3. A sphere of radius 0.5 Å is centered at the oxygen and at the two hydrogen atom probability maxima.
4. The 2×10^6 conformations of the Monte Carlo simulation are scanned to determine how many times a water molecule fell into the volume defined in step 3.

We note that in the probability density maps, the distance between the probability maxima of an oxygen atom and its associated hydrogen atoms is nearly (but not necessarily exactly) equal to the O—H distance in H_2O; thus, in our newly proposed algorithm, we lose some of the information available from the probability maps. In addition, the assumption of a sphere around the oxygen and hydrogen atoms, implies an isotropic probability distribution; as pointed out previously,[2-4] the probability distribution is often anisotropic, especially at room temperature. On the other hand, the advantage of the new method is that it allows us to obtain a graphical representation that is immediately understandable and replaces the snapshot pictures often used in the analysis of Monte Carlo data. As known, such pictures are limited to only *one* conformation and therefore have no statistical value. The technique described here brings 50–70% of the full set of the simulated conformations into the three spherical volumes associated with each water molecule. By increasing the sphere's radii to 1 Å, 90–95% of the computed configurations are accounted for; however, this larger radius decreased the information content concerning the relative orientation of the molecules in the solvent.

STRUCTURE OF FIRST SOLVATION SHELL

Let us start with a gross analysis of the computed data. In Fig. 1 we report the distribution of the internuclear distances from an atom (either O or H) of water to the nearest atom of the B-DNA fragment. The internuclear distance is designated as $R(i\text{-}a)$, where i refers either to a hydrogen or the oxygen atom in water and a refers to the atom of B-DNA nearest to i. We indicate as $N(i)$ the number of atoms of type i located (from analyses of the Monte Carlo Data) at a $R(i\text{-}a)$ distance from a. The (O-a) internuclear distance distribution (see the solid line of Fig. 1) shows a very distinct peak with a maximum at 2.6 Å (first region), a second peak with a maximum at 3.4 Å (second region), and a set of peaks that can be approximately associated to an asymmetrical Gaussian distribution with a maximum at about 4.5 Å (third region). The first, second, and third regions contain approximately 180, 40, and 220 water molecules, respectively.

The (H-a) internuclear distance distribution (broken line in Fig. 1) is characterized by two very distinct peaks with maxima at 1.8 and 2.8 Å and a distribution of peaks with decreasing intensity for large R(H-a) values. From these distributions we learn that *most* of the water molecules in the first region are oriented with a hydrogen atom pointing to a; since the (H-a) peak with a maximum at 1.8 Å is somewhat broader and a bit more intense than the second one (at 2.8 Å), we can also conclude that a small number of water molecules have *both* hydrogen atoms pointing towards B-DNA. The hydrogen and oxygen peak separations inform us (2.6 −1.8 Å = 0.8 Å and 2.6 − 2.8 Å = 0.2 Å) that, in general, for the first region the hydrogen bond forms structures for the type a- - -H—O.

A detailed analysis of the oxygen atom distribution in the first region indicates that relatively few water molecules have the oxygen atom (rather than a hydrogen atom) in proximity to an a atom. In general, we can associate the water molecules of the first region to the first hydration shell. The water molecules of the second R(O-a) region are less easily defined and might be interpreted as water molecules of the first shell but with the oxygen farther away from a and with both hydrogen atoms pointing toward

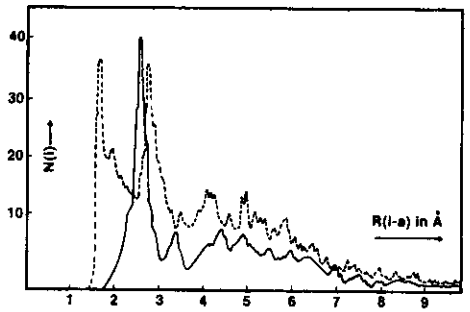


Fig. 1. Distribution of oxygen (solid line) and hydrogen (dashed line) internuclear distances from the a atom of B-DNA (see text).

a. [Namely, we suggest associating the second oxygen peak with the secondary peak of the hydrogen distribution at $R(i\text{-}a)$ of about 2.8 Å.]

Let us now analyze the water molecule distribution in more detail by considering the probability distribution maps, represented as explained in the Introduction. In Figs. 2–4 we show the water molecules solvating either the phosphate groups in each of the two helixes of B-DNA or the water molecules contained in a disk of about 5 Å thickness (slicing the cylinder perpendicular to the z-axis). The phosphates, the sugar groups, and the bases connected to one of the two helixes are designated with an asterisk in order to differentiate them from the equivalent groups of the second helix; equivalently, the two helixes are referred as h and h*, for short. In the figures the water molecules are represented with the new algorithm; we shall use the convenient expression of "water molecule number N" to indicate "the ensemble of water molecules that falls within the volume number N, consisting of the previously described three spheres of radius 0.5 Å.

In Fig. 2, we report the 10 phosphate groups P1–P10 of h and the corresponding sugar unit, but not the base pairs, which are, however, indicated

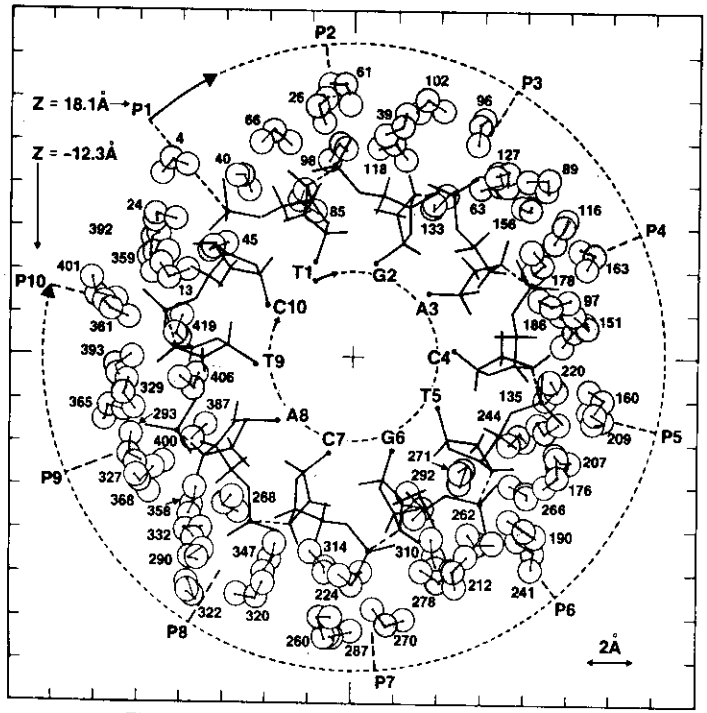


Fig. 2. Water molecules solvating the h helix in B-DNA.

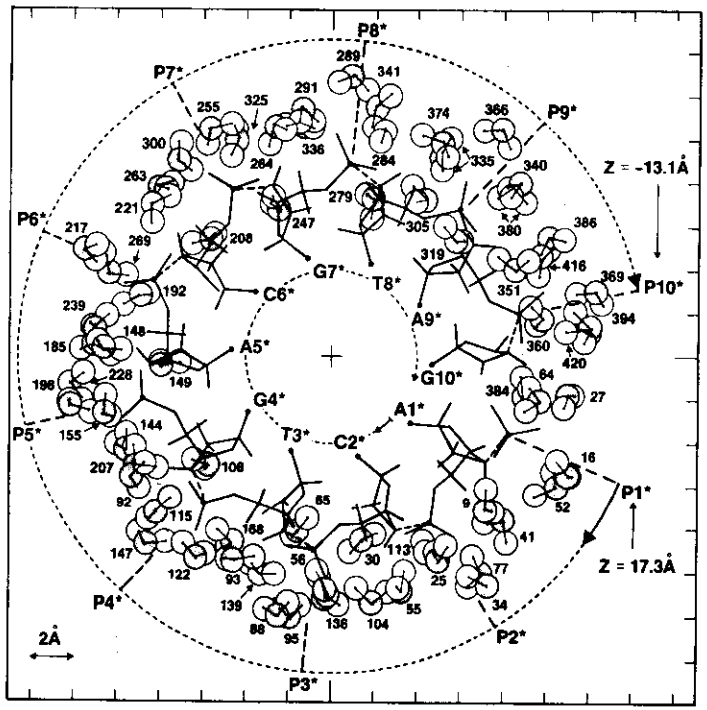


Fig. 3. Water molecules solvating the h* helix in B-DNA.

by reporting the terminal nitrogen atom and by using the notation, T1, G2, . . . , C10. The outermost circumference has 14.5-Å radius; the divisions around the figure indicate 2-Å intervals. The water molecules are seen from positive values looking down toward negative z values. Only the water molecules very *near* the phosphate P1 (at z = 18.1 Å) to P10 (at z = −12.3 Å) are reported. Since the "water molecules" are numbered with an index of increasing value along the z direction (from positive to negative z), low indices (starting from 1) correspond to water molecules solvating the top of the B-DNA fragment, high ones (approaching 447) to water molecules solvating the bottom of the B-DNA fragment.

Notice in Fig. 2 the dual features of the water clusters: not only do the water molecules enclose the PO_4^- group, they also form hydrogen-bonded filaments (see, for example, water molecules 322, 230, 332, and 358 in the P8–P9 region). The data in Fig. 3 provide a view of the water molecules solvating the phosphate groups of the h* helix. The water molecules in this figure seem to spread somewhat less along radial lines and to be confined to circular patterns. Close analysis (provided by the data in the tables that follow) mitigate this first-hand impression due to (a) steric visual ef-

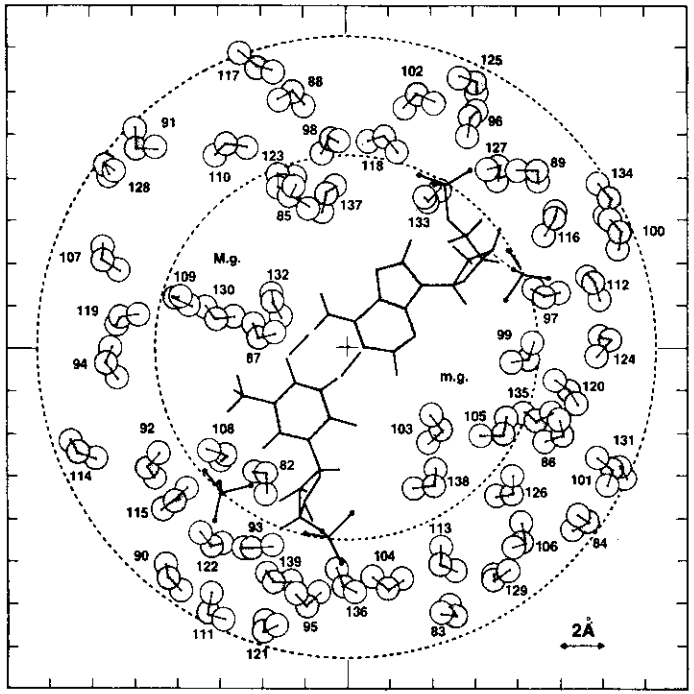


Fig. 4. Water molecules contained in a disk 4.4 Å thick and solvating B-DNA; the radii of the two circumferences are 14.5 and 8.8 Å, respectively.

fects and (b) an insufficient number of water molecules reported in the two figures. In Fig. 4, we report only one base pair, the A3·T3 base pair; the water molecules experience the immediate fields of the G2·C2* and C4·G4* base pairs (not reported in the figure in order not to further complicate the drawing).

Notice, in addition, that in the major groove (M.g.) there are 10 water molecules (82, 108, 87, 132, 130, 109, 123, 85, 137, and 133), whereas only 4 water molecules are found in the minor (m.g.) groove (138, 103, 105, and 99); this situation is reversed when we extend the major groove and the minor groove volume up to a radius of 14.5 Å. The physical reason for these findings is that close to the bases there is more "free" space in the major groove than in the minor groove, but further out toward $R = 13$ or 14 Å, the field generated by the phosphates in the minor groove is stronger than the field generated by the phosphates in the major groove.

A more quantitative analysis is provided in Tables I and II, where we report the water molecules solvating the phosphates of the h and h* helixes,

respectively. The first column reports the phosphate group identification; in the other columns, we report the water index, the distance from one of the hydrogen atoms (H1) from either a free oxygen of PO_4^- (namely, O1P and O2P) or the bound oxygen atoms (O3′ and O5′), the water–B-DNA average interaction energy (in kJ/mol), its mean standard deviation, and the water–water interaction energy (in kJ/mol) and its standard deviation.

In Figs. 5–8 we report the water molecules solvating the sugar unit and the base pairs. One point should be immediately stressed: we report in these figures not only the water molecules that are hydrogen-bonded to one (or more) bases, but also some of those "near" the bases. For a given water molecule, we use the following notation (in addition to the index for the water "volume"): *very strongly* hydrogen-bonded water molecules to B-DNA are differentiated from *strongly* hydrogen-bonded ones by writing a short identification of the solvated site either without or within parentheses.

For example, in Fig. 5 (1), the water molecules reported are in the vicinity of the T·A* and the G·C* base pair; the base pair (T·A*) is represented with solid lines, since it is at a higher z-value than the lower base pair (G·C*), represented by dashed lines. Water molecule 68 very strongly solvates the second phosphate group in the h helix (P2); water 46 very strongly solvates the zeroth sugar group in the h* helix (SO*); water 46 very strongly solvates SO* and also adenine A1* of h* and (but less strongly) guanine G2 of h; water 49 is too far from any B-DNA atom to be assigned as solvating a specific group, and it is therefore labeled only as M.g., namely, one of the molecules in the major groove; finally, water 73 is unlabeled, because it is neither strongly hydrogen-bonded to any atom of B-DNA nor is it within the major groove, and it therefore lies within the minor groove. In conclusion, we have considered three types of water molecules: *very strongly* hydrogen-bonded (to one or more atoms of B-DNA), *strongly* hydrogen-bonded, and *weakly* hydrogen-bonded. Somewhat arbitrarily, this classification is based on the value of the internuclear distance of the water's hydrogen (or oxygen) atoms from a given atom of the B-DNA fragment (see Fig. 1). More precisely, water molecules reported as *very strongly* hydrogen-bonded are those with an O-a internuclear distance not larger than 3.2 Å. If the hydrogen-bond length is larger by about 1 Å, then we classify it as a *weak* hydrogen bond. These criteria are rather restrictive and somewhat arbitrary: the PO_4^- field is very intense and a water molecule "strongly" hydrogen-bonded to one of the oxygen atoms in PO_4, necessarily *strongly* feels the field of the remaining atoms in the PO_4 group. For this reason, we have not even attempted to list in the tables those water molecules weakly or very weakly hydrogen-bonded to PO_4^-. In our classification of "hydrogen bond," we have included, as an additional criterion, the requirement that the overall orientation of a water molecular must be intuitively reasonable; for example, the oxygen (of H_2O) to oxygen (of PO_4^-) internuclear distance must be larger than the hydrogen (of H_2O) to oxygen (of PO_4) distance.

TABLE I
Water Molecules Solvating the Phosphates in the h Helix

P No.	W No.	Atom (1)	R(H1)	Atom (2)	R(H2)	E(W–D NA)	E(W–W)
P1	4	O1P	1.7	O1P	2.8	−115.3 ± 6.1	1.8 ± 3.5
	13	O2P	2.7	O2P	2.1	−86.9 ± 6.5	1.5 ± 1.2
	24	O2P	3.0	O2P	1.8	−114.5 ± 5.0	0.5 ± 3.8
	40	O1P	1.9	O5′	3.1	−90.0 ± 6.2	−12.0 ± 4.0
	45	O2P	1.5	O2P	2.9	−92.7 ± 3.7	−9.3 ± 4.2
P2	26	O1P	2.6	O3′	2.1	−103.9 ± 4.2	4.6 ± 4.5
	39	O1P	2.9	O1P	1.7	−104.8 ± 5.7	−2.2 ± 1.6
	61	O1P	1.6	O1P	2.9	−121.2 ± 5.7	3.9 ± 3.7
	66	O2P	1.7	O2P	2.8	−113.9 ± 4.8	0.8 ± 4.5
	68	O2P	2.6	O2P	3.0	−59.3 ± 3.8	−17.4 ± 4.0
	85	O2P	1.6	O2P	2.9	−104.1 ± 3.5	−7.5 ± 3.8
	98	O1P	2.4	O2P	3.0	−89.3 ± 3.9	−8.8 ± 4.8
P3	63	O3′	1.9	O3′	3.2	−60.4 ± 8.9	−18.3 ± 1.6
	89	O1P	2.4	O1P	3.2	−99.0 ± 5.8	−4.0 ± 3.5
	96	O1P	1.6	O1P	2.7	−113.7 ± 4.8	−10.4 ± 3.8
	118	O2P	1.6	O2P	2.9	−117.7 ± 5.6	−12.8 ± 3.3
	127	O1P	1.7	O1P	3.1	−124.5 ± 4.8	7.6 ± 3.5
	133	O2P	1.8	O2P	3.0	−70.6 ± 5.9	−21.7 ± 3.5
P4	97	O3′	1.8	O3′	2.9	−75.2 ± 4.6	−22.9 ± 2.6
	112	O1P	2.2	O1P	3.2	−81.5 ± 9.2	−15.2 ± 3.4
	116	O1P	3.3	O3′	2.1	−87.8 ± 6.6	0.3 ± 3.5
	151	O1P	1.7	O1P	2.8	−110.3 ± 6.7	−1.0 ± 4.2
	156	O2P	1.5	O2P	2.8	−119.7 ± 4.3	−10.8 ± 4.3
	163	O1P	1.7	O1P	2.8	−125.1 ± 6.8	2.3 ± 4.5
	178	O2P	2.1	O2P	3.0	−82.4 ± 3.6	−9.5 ± 3.5
P5	135	O3′	1.8	O3′	2.9	−61.9 ± 3.8	−18.2 ± 4.3
	160	O1P	2.0	O1P	2.9	−104.2 ± 5.0	2.0 ± 5.3
	176	O1P	1.7	O1P	3.1	−117.9 ± 7.2	2.9 ± 4.5
	186	O2P	1.6	O2P	2.8	−121.5 ± 3.6	−13.7 ± 3.2
	209	O1P	1.7	O1P	2.8	−130.8 ± 4.8	−0.2 ± 3.5
P6	190	O3′	1.8	O3′	3.2	−80.3 ± 5.9	−12.2 ± 4.7
	212	O1P	1.6	O1P	2.9	−85.4 ± 7.5	−10.5 ± 4.0
	241	O1P	1.6	O1P	2.8	−105.8 ± 5.6	−8.9 ± 4.3
	244	O2P	1.6	O2P	2.7	−91.1 ± 3.5	−10.0 ± 4.0
	262	O1P	2.0	O1P	2.8	−77.3 ± 7.2	−14.6 ± 3.0
	266	O2P	2.1	O2P	2.6	−144.3 ± 6.1	2.1 ± 5.4
	271	O2P	2.3	O2P	3.5	−75.7 ± 7.0	−18.0 ± 5.6
P7	224	O3′	1.7	O3′	3.1	−95.3 ± 4.3	−6.9 ± 2.1
	260	O1P	1.9	O1P	2.7	−105.0 ± 6.2	−0.1 ± 5.0
	270	O1P	1.7	O1P	2.9	−125.6 ± 4.5	7.1 ± 5.1
	278	O2P	1.7	O2P	2.9	−128.3 ± 4.6	−4.8 ± 3.4
	292	O2P	1.6	O2P	2.9	−104.9 ± 5.1	−2.1 ± 5.0
P8	268	O3′	1.8	O3′	3.1	−90.3 ± 5.2	−13.9 ± 3.9
	290	O1P	1.6	O1P	3.0	−121.2 ± 5.8	2.9 ± 3.2
	314	O2P	1.7	O2P	3.0	−113.5 ± 11.2	3.0 ± 2.9
	320	O2P	2.5	O1P	2.8	−93.1 ± 9.8	−9.2 ± 3.6
	322	O1P	2.9	O1P	2.9	−54.0 ± 10.7	−17.8 ± 3.9

(*continued*)

TABLE I (*continued*)

P No.	W No.	Atom (1)	R(H1)	Atom (2)	R(H2)	E(W–D NA)	E(W–W)
	332	O1P	1.7	O1P	2.9	−108.5 ± 5.4	−10.6 ± 2.9
	347	O2P	1.8	O2P	2.7	−114.0 ± 3.4	−11.3 ± 4.2
P9	327	O3′	2.0	O3′	3.1	−76.3 ± 6.6	−9.1 ± 5.1
	329	O1P	1.9	O1P	2.8	−93.2 ± 3.4	1.0 ± 3.9
	358	O2P	1.5	O2P	2.8	−125.3 ± 4.7	−5.7 ± 4.2
	365	O1P	1.6	O1P	2.8	−94.7 ± 5.6	−7.4 ± 3.2
	368	O1P	1.9	O2P	2.7	−119.3 ± 10.6	8.5 ± 4.4
	387	O1P	1.8	O2P	3.3	−109.2 ± 7.4	−4.1 ± 3.1
P10	359	O3′	2.4	O3′	3.5	−86.6 ± 4.1	−5.7 ± 3.4
	361	O3′	1.9	O1P	2.5	−85.3 ± 6.9	8.4 ± 4.2
	392	O1P	1.6	O1P	2.8	−109.7 ± 5.0	−5.3 ± 2.8
	393	O2P	1.7	O2P	3.0	−124.3 ± 4.5	3.1 ± 4.7
	401	O1P	1.6	O2P	2.6	−116.7 ± 3.0	−5.5 ± 3.8
	406	O2P	2.0	O2P	2.8	−83.8 ± 7.9	−3.2 ± 5.6
	419	O2P	1.7	O2P	3.1	−109.4 ± 7.0	−0.8 ± 2.6
P11	408	O1P	1.6	O1P	2.9	−113.0 ± 5.1	−5.7 ± 5.0
	432	O2P	1.8	O2P	2.8	−79.3 ± 4.9	−18.0 ± 5.6
	433	O1P	1.6	O1P	2.7	−118.3 ± 4.3	8.5 ± 2.5
	439	O2P	1.8	O2P	2.8	−120.4 ± 4.3	−2.5 ± 4.4

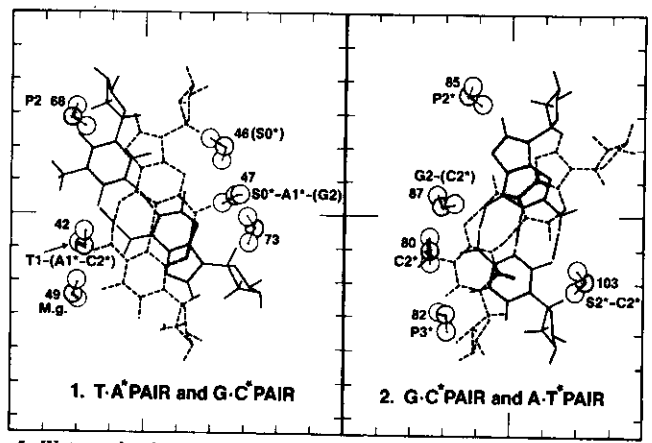


Fig. 5. Water molecules in the vicinity of T1·A1* and G2·C2* (1) and G2·C2* and A3·T3* (2).

In Tables III and IV we present the water molecules hydrogen-bonded to the sugar units and to the bases, respectively. The absence of solvating water molecules at some of the sugar units (or bases) is a consequence of the above restrictive definitions. By relaxing the hydrogen-bond length criteria, more molecules would be assigned to a given group of B-DNA.

TABLE II
Water Molecules Solvating the Phosphates in the h* Helix

P No.	W No.	Atom (1)	R (H1)	Atom (2)	R (H2)	E(W–DNA)	E(W–W)
P1*	9	O2P	1.6	O2P	2.8	−116.2 ± 5.5	−5.2 ± 3.5
	16	O1P	1.7	O1P	2.8	−114.9 ± 4.1	−2.8 ± 3.1
	27	O1P	1.6	O1P	2.8	−114.3 ± 3.9	−3.5 ± 4.0
	64	O3′	2.2	O3′	3.6	−76.1 ± 4.1	−16.6 ± 3.7
P2*	25	O1P	3.1	O2P	2.0	−109.7 ± 5.3	−3.5 ± 4.4
	30	O2P	1.7	O2P	2.9	−115.3 ± 6.7	3.2 ± 1.6
	34	O1P	2.7	O1P	2.8	−77.8 ± 6.1	−13.8 ± 3.5
	41	O1P	1.7	O2P	2.6	−83.9 ± 4.8	−16.4 ± 4.2
	55	O1P	2.0	O1P	2.6	−86.4 ± 6.0	−7.3 ± 3.1
	77	O1P	1.6	O1P	2.9	−104.6 ± 6.5	−9.4 ± 3.9
P3*	65	O2P	1.7	O2P	3.0	−114.5 ± 4.2	−3.0 ± 2.6
	93	O2P	1.7	O2P	3.0	−120.2 ± 6.7	−7.8 ± 3.5
	95	O1P	1.7	O1P	2.9	−120.7 ± 6.0	−1.5 ± 5.6
	104	O1P	1.6	O1P	3.0	−114.8 ± 6.7	−2.9 ± 4.7
	136	O1P	2.7	O3′	1.9	−85.6 ± 5.2	−7.5 ± 4.6
P4*	108	O2P	1.7	O2P	2.8	−80.4 ± 6.0	−14.3 ± 5.2
	115	O2P	1.8	O2P	2.8	−112.3 ± 6.7	−4.6 ± 3.6
	122	O1P	1.7	O1P	2.4	−106.2 ± 7.4	−7.8 ± 3.5
	144	O2P	1.7	O2P	3.1	−111.4 ± 8.5	−8.4 ± 2.5
	147	O1P	1.7	O1P	3.0	−119.4 ± 7.2	6.9 ± 4.5
	168	O1P	1.9	O1P	3.4	−74.5 ± 13.2	−11.2 ± 3.3
P5*	148	O2P	2.8	O2P	3.0	−84.5 ± 6.3	−10.6 ± 5.4
	149	O2P	1.7	O2P	3.1	−115.3 ± 7.1	0.0 ± 2.9
	155	O1P	1.8	O2P	2.9	−121.1 ± 8.2	1.3 ± 5.0
	185	O2P	1.7	O2P	2.9	−106.4 ± 6.9	−12.5 ± 3.0
	198	O1P	1.6	O1P	3.0	−120.5 ± 6.8	0.8 ± 3.6
	207	O1P	1.6	O1P	2.8	−92.3 ± 4.8	−5.9 ± 4.2
P6*	192	O1P	3.0	O5′	2.5	−75.4 ± 5.9	−11.7 ± 4.4
	208	O2P	1.8	O2P	2.5	−94.3 ± 7.0	−14.8 ± 3.3
	217	O1P	1.7	O1P	2.7	130.7 ± 4.0	1.8 ± 3.4
	221	O2P	1.7	O2P	2.8	−127.5 ± 4.2	−1.2 ± 3.6
	239	O1P	1.6	O2P	2.7	−97.8 ± 6.3	−8.1 ± 2.7
	269	O3′	1.8	O3′	3.0	−84.2 ± 4.9	−19.9 ± 4.5
P7*	237	O1P	2.6	O1P	3.3	−54.9 ± 4.8	−22.6 ± 5.1
	247	O2P	1.6	O2P	2.8	−82.1 ± 4.1	−17.8 ± 2.9
	255	O1P	1.7	O1P	2.9	−127.9 ± 5.8	1.0 ± 4.7
	263	O1P	1.6	O1P	2.8	−118.4 ± 6.9	8.1 ± 4.4
	264	O2P	1.6	O2P	2.8	−119.2 ± 4.3	−7.9 ± 3.7
	300	O1P	1.8	O1P	2.7	−95.4 ± 4.6	−5.5 ± 4.2
	325	O3′	2.4	O3′	3.5	−41.7 ± 7.7	−25.4 ± 3.6
P8*	279	O2P	2.0	O2P	3.5	−96.6 ± 5.7	−15.7 ± 3.1
	284	O2P	1.7	O2P	2.8	−128.6 ± 2.5	−4.1 ± 3.5
	291	O1P	1.7	O1P	3.1	−126.7 ± 7.9	−0.1 ± 5.4
	395	O2P	1.8	O2P	2.6	−95.2 ± 9.0	−8.9 ± 2.8
	326	O1P	2.8	O1P	3.0	−56.1 ± 10.6	−19.2 ± 3.6
	337	O1P	1.7	O1P	2.8	−118.0 ± 4.5	−3.1 ± 2.6
	341	O1P	2.7	O3′	2.0	−95.3 ± 5.4	−12.9 ± 4.4

(continued)

TABLE II (*continued*)

P No.	W No.	Atom (1)	R (H1)	Atom (2)	R (H2)	E(W–DNA)	E(W–W)
P9*	319	O2P	1.7	O2P	3.1	−89.9 ± 5.3	−17.3 ± 3.1
	335	O1P	1.7	O1P	2.8	−65.3 ± 8.0	−16.9 ± 4.2
	340	O1P	2.1	O2P	3.0	−117.4 ± 4.4	2.0 ± 4.2
	351	O2P	1.6	O2P	2.7	−88.0 ± 5.2	−17.8 ± 2.5
	366	O1P	2.7	O1P	2.8	−88.7 ± 6.5	−0.2 ± 4.4
	374	O1P	1.6	O1P	3.1	−113.1 ± 8.2	−6.2 ± 5.1
	380	O3′	1.9	O3′	3.1	−90.9 ± 4.7	−9.3 ± 3.7
P10*	360	O2P	1.8	O2P	3.1	−103.5 ± 5.0	−10.6 ± 3.3
	384	O2P	1.6	O2P	2.9	−106.6 ± 5.3	−9.8 ± 4.2
	386	O1P	1.8	O1P	2.9	−89.1 ± 8.0	−20.1 ± 3.2
	394	O1P	1.8	O2P	2.9	−125.2 ± 4.2	0.3 ± 2.7
	416	O1P	2.2	O1P	2.4	−106.4 ± 5.5	−1.9 ± 3.1
P11*	399	O2P	1.7	O2P	3.1	−93.5 ± 6.7	−3.3 ± 5.0
	402	O1P	1.7	O1P	3.0	−115.6 ± 4.0	−5.6 ± 3.2
	409	O2P	2.2	O2P	2.8	−96.8 ± 6.4	−7.4 ± 2.8
	421	O2P	1.7	O2P	2.8	−119.1 ± 6.4	3.8 ± 5.0
	435	O1P	2.1	O1P	2.6	−106.4 ± 4.5	−2.3 ± 5.0
	441	O1P	1.8	O1P	2.6	−105.8 ± 5.8	3.3 ± 2.8

AVERAGE INTERACTION ENERGIES

With these restrictive definitions and considering only *strongly* hydrogen-bonded water molecules, we can summarize as follows: there are 5.9 water molecules solvating each PO_4 group, 0.3 water molecules solvating each sugar group, 0.5 water molecules solvating both the sugars and bases (namely, hydrogen-bonded bridges between a sugar and a base), and 0.9 water molecules solvating a base. These average values refers only to the first solvation shell, in the strict sense above defined; in this way, about 160 water molecules out of a total of 447 are considered. Considering the

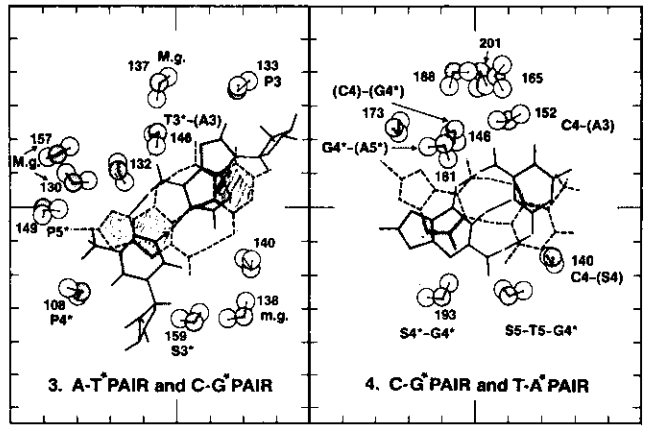


Fig. 6. Water molecules in the vicinity of A3·T3* and C4·G4* (3) and C4-G4* and T5-A5* (4).

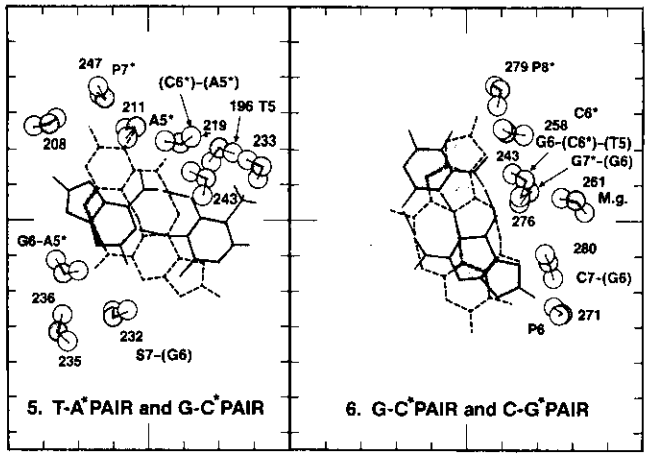


Fig. 7. Water molecules in the vicinity of T5·A5* and G6·C6* (5) and G6·C6* and C7·G7* (6).

grooves, there are 83 and 103 water molecules in the first solvation shell of the minor and major grooves, respectively; outside the first solvation shell there are an additional 83 and 177 water molecules in the minor and major grooves, respectively. The average water–B-DNA interaction energies (in kJ/mol) are -101.9 ± 5.8, -86.9 ± 4.3, -85.9 ± 4.4, and -63.4 ± 4.2 for the PO_4^-, sugar, sugar and base, bases, respectively; the average water–water interaction energies are -6.1 ± 3.8, -12.6 ± 3.7, -12.6 ± 3.5, and -16.6 ± 4.3 kJ/mol for the same groups listed above. Let us now consider the energetics of the full set of 447 water molecules (see Fig. 9).

The interaction energies (water–water, water–B-DNA, and total) are

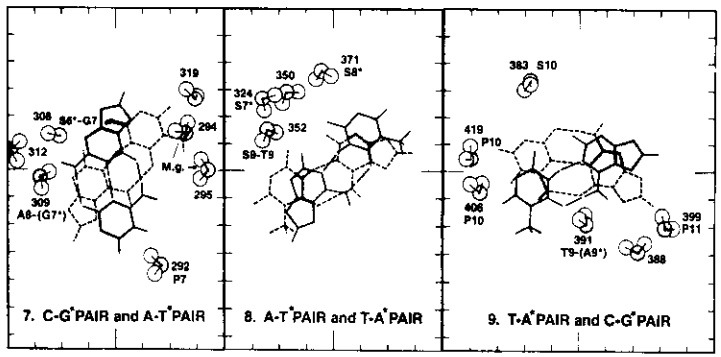


Fig. 8. Water molecules in the vicinity of C7·G7* and A8·T8* (7), A8·T8* and T9·A9* (8), and T9·A9* and C10·10* (9).

TABLE III
Water Molecules Solvating the Sugar Units

S No.	W No.	R (H1)	R (H2)	E(W–DNA)	E(W–W)	Notes
S2	46	1.9	3.3	-98.1 ± 4.6	-9.2 ± 4.7	S2
S5	179	1.9	3.1	-106.2 ± 4.3	-10.6 ± 2.6	S5-T5-G4*
S7	232	2.2	3.7	-90.8 ± 2.9	-11.5 ± 2.6	S7-(G6)
S8	309	1.7	3.0	-89.5 ± 5.5	-6.2 ± 4.2	S8-A8-(G7*)
S9	352	1.8	3.0	-84.3 ± 4.1	-15.1 ± 2.5	S9-T9
S10	383	1.9	2.9	-78.7 ± 4.2	-16.9 ± 3.2	S10
S0*	47	2.0	3.2	-48.4 ± 6.9	-23.3 ± 3.5	S0*-A1*-(G2)
S2*	103	1.8	2.9	-89.4 ± 3.9	-16.0 ± 4.7	S2*-C2*
S3*	159	1.7	3.0	-84.8 ± 3.1	-19.3 ± 4.8	S3*
S4*	193	2.0	3.0	-81.9 ± 5.7	-14.4 ± 4.4	S4*-G4*
S5*	236	1.9	3.1	-100.5 ± 2.9	-9.4 ± 3.9	S5*-G6-A5*
S6*	308	2.0	3.4	-92.0 ± 3.3	-7.2 ± 2.8	S6*-G7
S7*	324	1.8	2.9	-89.8 ± 3.8	-4.8 ± 2.8	S7*
S8*	371	2.0	3.3	-83.6 ± 5.2	-12.3 ± 5.1	S8*

reported for the water molecule with its oxygen atom enclosed in the interface of two coaxial cylinders differing in the two radii by 0.5 Å. The first point corresponds to the energy of the water molecules enclosed in a cylinder of radius $R = 3.5$ Å (it contains no water molecules). Extrapolation from $R = 14.5$ Å to very large R qualitatively is easy, since the water–DNA interaction will go to zero and the water–water interaction will go to the simulated value for bulk water (35.6 kJ/mol; see Refs. 10, 11). Quantitatively, however, the extrapolation is somewhat more difficult, and we would like to limit ourselves to present the simulated data; namely, for a water molecule from $R = 0$ to 14.5 Å, the average total energy is -79.5 ± 0.1 kJ/mol, and the average water–water interaction energy is -16.3 ± 0.05 kJ/mol; the simulated value for total energy is an upper limit to the water–DNA interaction energy. Extension of our simulation to about 1500–2000 water molecules would fully settle this point, but since the necessary simulation is, presently, somewhat too expensive, we have deferred it to a later data. The total energies above given can be somewhat misleading, since it is obtained as an average over many different positions, with large energy variations. The total interaction energy of Fig. 9 clearly shows five minima corresponding to interactions with the bases (A), bases and sugars (B), and phosphates (C, D, and E). The O3′ and O5′ solvation region and the O1P and O2P solvation region approximately correspond to the C and E minima, respectively; the minimum D corresponds to the solvation region intermediate between C and E. The energy curve is computed up to 14.5 Å and "freely" extrapolated thereafter. The water–water (W-W) and water–DNA (W-BDNA) components of the interaction energy support the above interpretation for the energy-minima pattern; however, the partial interaction energies are more structured in the region from 8 to 14.5 Å; the three to four minima seem to provide a differentiation between the four oxygen atoms of PO_4. This analysis, however, is com-

TABLE IV
Water Molecules Solvating the Base Pairs

W No.	Base	R (1)	Atom (1)	R (2)	Atom (2)	E(W–DNA)	E(W–W)	Notes
42	T1	1.9	O4			−70.5 ± 5.8	−12.0 ± 3.5	T1-(A1*)-(C2*)
	A1*	2.6	HN61					
	C2*	3.0	HN42					
47	A1*	2.2	N3			−48.4 ± 6.9	−23.3 ± 3.5	SO*-A1*-(G2)
	G2	2.8	HN21					
80	C2*	1.9	HN42			−65.5 ± 2.9	−15.8 ± 4.7	C2*
87	G2	1.9	O6			−75.5 ± 5.1	−7.5 ± 6.1	G2-(C2*)
	C2*	2.8	HN41	3.0	HN42			
103	C2*	2.0	O2	3.0	O2	−89.4 ± 3.9	−16.0 ± 4.7	S2*-C2*
132	T3*	2.0	O4			−75.6 ± 4.3	−6.3 ± 5.2	T3*-(A3)
	A3	2.6	HN61	2.6	HN62			
140	C4	1.9	O2	3.0	O2	−81.9 ± 3.4	−13.8 ± 4.3	C4-(S4)
146	C4	2.6	HN42			−61.0 ± 3.3	−17.8 ± 4.7	(C4)-(G4*)
	G4*	2.9	O6					
152	C4	1.9	NH42			−55.0 ± 2.8	−26.5 ± 3.9	C4-(A3)
	A3	2.7	N7					
179	T5	2.0	O2	2.9	O2	−106.2 ± 4.3	−10.6 ± 2.6	S5-T5-G4*
	G4*	2.3	HN21					
181	G4*	1.9	O6			−79.9 ± 4.7	−18.0 ± 3.3	G4*-(A5*)
	A5*	2.9	N7	2.7	HN61			
193	G4*	2.3	N3			−81.9 ± 5.7	−14.4 ± 4.4	S4*-G4*
196	T5	2.3	O4			−55.2 ± 4.1	−12.3 ± 4.4	T5
211	A5*	2.3	HN61			−60.6 ± 3.3	−27.3 ± 5.1	A5*
219	A5*	3.0	HN61			−26.0 ± 7.8	−33.9 ± 5.0	(C6*)-(A5*)
	C6*	2.8	HN42					
232	G6	2.5	N3	2.6	N3	−90.8 ± 2.9	−11.5 ± 2.6	S7-(G6)
236	G6	2.3	HN21			−100.5 ± 2.9	−9.4 ± 3.9	S5*-G6-A5*
	A5*	2.3	N3					
243	G6	2.0	O6	2.8	O6	−69.9 ± 4.3	−10.7 ± 2.5	G6-(C6*)-T5
	C6*	2.8	HN42					
	T5	2.8	O4	2.8	O4			
258	C6*	1.9	NH42			−26.0 ± 3.0	−31.7 ± 5.0	C6*
276	G7*	1.9	O6			−69.4 ± 3.7	−16.1 ± 2.8	G7*-(G6)
	G6	2.5	O6	2.9	O6			
280	C7	1.8	HN42			−77.2 ± 3.7	−14.6 ± 2.8	C7-G6
	G6	2.4	N7					
308	G7*	2.8	N3	2.3	N3	−92.0 ± 3.3	−7.2 ± 2.8	S6*-G7
309	A8	2.2	N3			−89.5 ± 5.5	−6.2 ± 4.2	S8-A8-(G7*)
	G7*	2.4	HN21					
352	T9	2.0	O2	2.7	O2	−84.3 ± 4.1	−15.1 ± 2.5	S9-T9
391	T9	1.9	O4			−83.4 ± 3.9	0.9 ± 5.0	T9-(A9*)
	A9*	2.8	HN61					

plicated by the fact that much water is located in the major and the minor grooves. Such water has little resemblance to bulk water, since it is highly structured, and therefore, its energy is not a "constant additive" contribution to the energy curve.

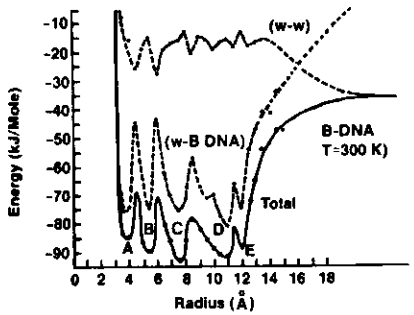


Fig. 9. Interaction energies of a water molecule with B-DNA in kJ/mol.

TRANSGROOVE AND INTERPHOSPHATE WATER FILAMENTS

The complexity of the structure of water in the two grooves is evident in Fig. 10, where we consider the water molecules enclosed in the volume between two coaxial cylinders of radii 8 and 12 Å, respectively; the figure reports the water of only one-half of the cylinder volume on a y,z projection. To us the striking feature of this figure is clear evidence of hydrogen-bonded water filaments from a phosphate group of h to a phosphate group of h*, spanning the major groove (the filaments have nearly periodic rototranslational symmetry) and of hydrogen-bonded water filaments connecting two successive phosphate groups, from P(i) and P(i + 1) in the h (and/or h*) helix. We have previously commented on these *filaments*,[2–5] since they are either implicit in the isoenergy maps or explicit in the A-DNA and B-DNA single-helix Monte Carlo simulations; however, the limited number of water molecules considered in our previous simulations limited the validity of our suggestion. We note that this feature—the existence of filaments—has been previously encountered in ion pairs in solution.[12] We feel that this feature is basic in any water solution containing ions and will have profound consequences to the understanding of dynamical and temperature-dependent properties of solutions with ions.

The hydrogen bonds shown in Fig. 10 were reported only if the oxygen–oxygen distance (between two waters) is equal or smaller than 3.5 Å, and if the oxygen–hydrogen distance is smaller than the corresponding oxygen–oxygen distance. Typical water filaments (Fig. 10) are formed by the waters 310, 323, 343, 377, 388, and 399 linking P7 of h to P11* of h*; or 271, 266, 292, 304, 339, 357, 348, and 364 linking P6 of h to P10* of h*; or 220, 257, 288, 299, 328, and 340 linking P5 of h to P9* of h*. These are *transgroove* filaments. Other structured filaments are present and connect a phosphate to a successive phosphate in the same helix. Notice, for example, the water molecules for the *interphosphate* filaments 402, 384, 364, and 394 from P11* and P10* and 364, 360, and 351 from P10* to P9*. In

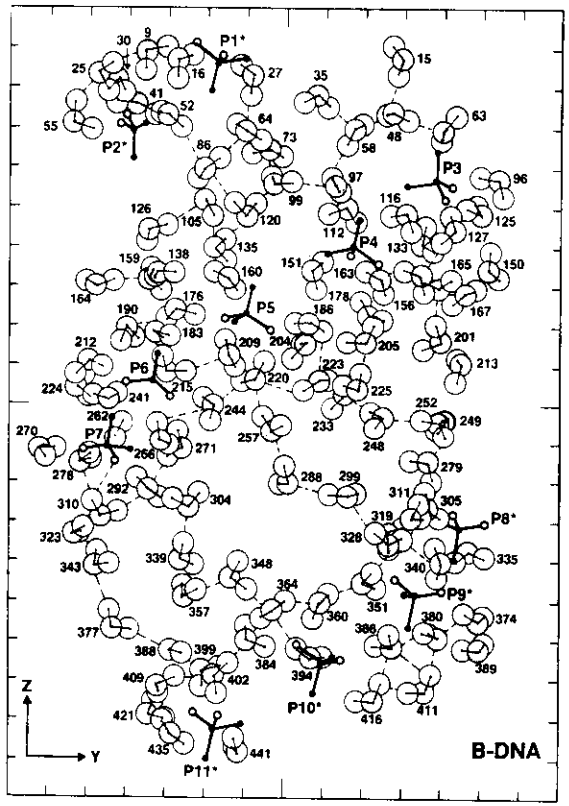


Fig. 10. Network of water molecules in B-DNA (see text).

the above examples, waters 310 and 399 are *terminal waters of a transgroove filament;* the structure in the *interphosphate* filaments is different, since filaments 402, 384, and 369 (from P11* to P10*) *continue* with waters 360 and 351, leading to P9*. Given that we are discussing a network complex, there is some element of arbitrariness in defining terminal waters in a fragment; however, the finding of two different structural organizations, namely, the *transgroove* filament and the *interphosphate* filaments, seems firm. It is stressed that these structures do not correspond to data obtained by analyzing *one* or *few* conformations, but are statistically "stable" and meaningful structures. It is very tempting to postulate that protons "are transferred preferentially along these filaments"; hence, these filaments are of importance in reactivity studies. These structures are "dynamical" in the sense that a given structure can evolve into a different structure, at relatively little expense for the *total* energy of the system.

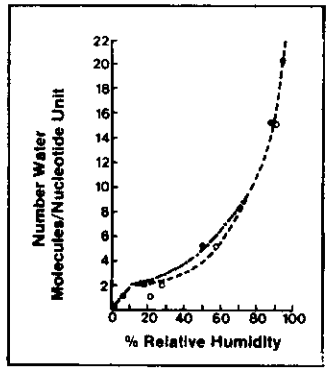


Fig. 11. Simulated Monte Carlo partial (○) and total (●) isotherms at 300 K for B-DNA and experimental Na^+-DNA desorption (- - -) and adsorption (— —) isotherms at 300 K.

ISOTHERM SIMULATION

Having briefly analyzed these microaspects of the DNA solvation, let us now turn to the other extreme and analyze some macroaspect, in particular, the absorption and desorption isotherms. These have been studied experimentally in depth.[13-24] We refer, in addition, to the short but recent review by Texter on this point.[25] The absorption–desorption hysteresis cannot be obtained directly from Monte Carlo simulations if the solute is not allowed to structurally adjust itself to concentration variations of the solvent. However, the hysteresis is a "fine detail" of the isotherm; the main characteristic of the isotherm is the well-known sigmoidal shape. We note that our simulation (447 water molecules) corresponds to 92% of relative humidity, or 20.3 water molecules per nucleotide unit.[13] If we start with 447 water molecules and subtract progressively an increasing number of water molecules (desorption simulation) *without* allowing the water molecules to rearrange, the simulated isotherm is not sigmoidal but nearly

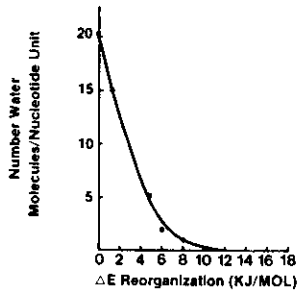


Fig. 12. Reorganization energy of water molecules relative to a sample with 92% water humidity (in B-DNA) at 300 K (see text).

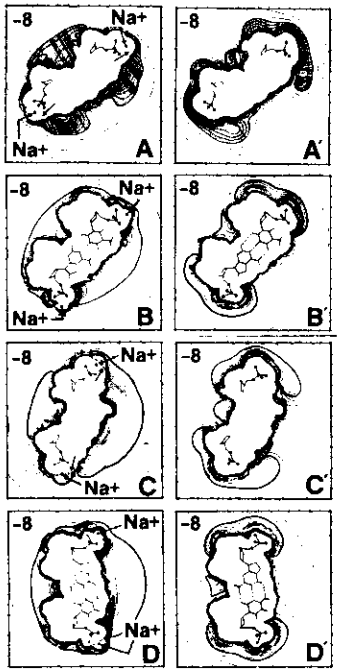


Fig. 13. Isoenergy maps for the interaction energy of one water molecule with B-DNA (right) and Na^+–B-DNA (left).

linear. If, however, we subtract water molecules *and reperform* Monte Carlo simulations, then the water molecules can rearrange themselves: in this case, the isotherm has a nicely sigmoidal shape (see Fig. 11). By construction (see above), the simulated desorption and absorption isotherms are equal. In Fig. 11 we have also reported the desorption and absorption isotherms for Na-DNA.[13] Clearly, the two experiments (our simulation and the Na-DNA isotherms) should not be compared in an attempt to find a one-to-one correspondence. In Fig. 12 we report the water reorganization energy for the water molecules solvating B-DNA, namely, the energy differences between Monte Carlo simulations with N water molecules (with $N < 447$) and simulations with 447 water molecules. It is evident from Fig. 11 that the reorganization energy increases by decreasing the number of molecules in the solvent, as clearly expected. In our new simulation we have selected the following values for N: 333, 114, 44, and 22. Lowering the relative humidity first brings about the elimination of weakly bound water (from large R values to $R = 13$ Å), then elimination of the *transgroove* filaments, then of water from the vicinity of the bases and from the sugar, and finally from the *interphosphate* filaments.

Removal and/or rearrangement of the transgroove filaments brings about a rearrangement of the B-DNA structure; that in turn induces a second water rearrangement proportional to the width of the hystersis loop. Therefore, simulation experiments of the type here reported *coupled* with laboratory isotherms experiments will allow differentiating between water rearrangement and DNA-induced rearrangement. Unfortunately, this comparison cannot be made with B-DNA without counterions.

B-DNA DOUBLE HELIX WITH NA^+ COUNTERIONS

Let us now briefly comment on some aspects of the solvation of the Na^+-DNA double helix, limiting ourselves to the qualitative findings based on the isoenergy maps reported in Fig. 13. The 22 phosphate groups of our B-DNA fragment have been neutralized by placing a Na^+ ion at the energy-minimum position of Na^+ interacting either with diethyl phosphate or other phosphates.[26,27] The four panels A–D on the *left* of Fig. 13 correspond to isoenergy maps for a water molecule interacting with our B-DNA fragment having one Na^+ ion at each phosphate (hereafter referred to as Na^+–B-DNA, for short). The four planes reported in the panels are mutually parallel and contain the phosphate groups above the A·T* pair (A), the A1·T1* pair (B), the phosphate group above the G2-C2* pair (C), and the G2·C2* pair (D). The isoenergy maps in the four panels to the *right* (A′–D′) are in corresponding planes for B-DNA. The contour corresponding to −8.0 kcal/mol (outermost contour) is explicitly indicated; the contour-to-contour energy difference is 2.0 kcal/mol. Comparing A and A′ in Fig. 13, we notice that the area delimited by the contour at −14.0 kcal/mol and by the hard core (the most attractive region, shaded areas in Fig. 13, panels A and A′) increases in the vicinity of the sugars and bases in A relative to A′. The same feature is evident by comparing B with B′, C with C′, and D with D′. In the maps of B, B′, C, C′, D, and D′, the contour at −14.0 kcal/mol is distinguished by heavier lines (relative to the remaining contours). As a consequence of the energy variations in the most attractive regions, we expect that the water molecules in the minor and major grooves will be more attracted in Na^+–B-DNA than in B-DNA; the same comment holds for the water solvating the base pairs. The *large* difference in the contour maps between B-DNA and Na^+–B-DNA ensures that there are corresponding differences in the water structure, namely, (a) in the number of water molecules solvating the first shell and (b) in the structure of the hydrogen-bonded filaments, previously discussed. We can restate this point in a different way. The type of reasoning presented to explain the hysteresis in the isotherms is here advanced to support the expectation of induced rearrangements in the DNA structure. However, the isoenergy contour differences between Na^+–B-DNA and B-DNA are *very* large; therefore, we expect *significant* variations in the DNA structure, in agreement with the experimentally well-known transitions from one conformation of DNA to another, induced by variations in the concentration and/or type of counterions.

CONCLUSIONS

It is known[25] that by increasing the relative humidity above ~90%, the added water molecules exhibit a bulk-water-type behavior. Therefore, our sample of 447 water molecules (corresponding to 92% relative humidity) is sufficiently extended to describe in a meaningful way not only the first solvation shell, but also the water molecules enclosed in the major and minor grooves. Additional water molecules would provide a better definition in the energy diagram (Fig. 9) at large R values; however, several hundred more water molecules would have to be added to describe only few angstroms above R = 14.5 Å.

The experimental sequence of hydration at different molecular sites[14] is reported to occur first at the two free oxygens atoms or PO_4^-, then at the two bonded oxygen atoms of PO_4^-, then at the oxygen atom of the sugar, and finally at the bases. These findings are corroborated both by our simulation and complemented by energy data on the water-site interaction energy. We find that the most attractive site is PO_4^-, followed by the sugar unit, and finally by the bases. By combining the information of Figs. 1 and 9, one can distinguish between the two free-oxygen atoms and the two bonded oxygen atoms in PO_4^-. However, one must keep in mind the fact that the PO_4^- field is very attractive for water molecules, and therefore, a sharp distinction in the attraction for a water molecule from one of the two types of oxygen atoms independently from the other is not too meaningful. On the other hand, we recall that the field near the two free oxygen atoms is reinforced by the field of the two bonded oxygen atoms, whereas the field near the two bonded oxygen atoms is reinforced by the field of the two free oxygen atoms but, at the same time, weakened by the field of the —CH_2 group.

The number of water molecules solvating the PO_4^- group at room temperature has been estimated[14] to be between 5 and 6 in DNA neutralized by Na^+ counterions; our simulation yields 5.9 water molecules for B-DNA double helix without counterions.

Indirect experimental evidence (angular distribution of near elastic scattering by neutron diffraction[28]) suggests the existence of transgroove water molecules (preferential orientation along the main DNA axis). In our simulation we have provided detailed information on the transgroove water molecules, and we have shown that these form water molecule filaments, hydrogen-bonded, connecting a phosphate group at one side of the major groove with another phosphate group at the opposite site of the major groove. In addition to the transgroove water molecule filaments, there are interphosphate filaments, shorter in length than those of the transgroove one. The highly ordered texture of water molecules is intimately connected with the B-DNA structure and stability and will clearly change for conformational transitions in B-DNA. (For additional discussions, see Ref. 29.)

Finally, by providing the isotherm of B-DNA double helix we have opened the way to the possibility of comparing isotherms in B-DNA either

with or without counterions. In our opinion, such comparisons are essential in order to differentiate between the "water-reorganization" contribution and the counterion contribution to conformational transitions.[30] This differentiation is of importance for any detailed explanation of the mechanisms associated with the absorption–desorption hystereses. In a paper to follow, we shall present a detailed analysis of water molecules solvating B-DNA with Na^+ counterions and a very detailed comparison with the water organizations at different relative humidities.[30]

References

1. Scordamaglia, R., Cavallone, F. & Clementi, E. (1977) *J. Am. Chem. Soc.* **99,** 5545–5549.
2. Clementi, E. & Corongiu, G. (1980) *J. Chem. Phys.* **72,** 3979–3992.
3. Clementi, E. & Corongiu, G. (1979) *Biopolymers* **18,** 2431–2450.
4. Clementi, E. & Corongiu, G. (1979) *Int. J. Quant. Chem.* **16,** 897–915.
5. Clementi, E. & Corongiu, G. (1979) *Chem. Phys. Lett.* **60,** 175–178.
6. Clementi, E. & Corongiu, G. (1979) *Gazz. Chim. Ital.* **109,** 201–205.
7. Romano, S. & Clementi, E. (1980) *Int. J. Quant. Chem.* **17,** 1007–1029.
8. Clementi, E., Corongiu, G., Jonsson, B. & Romano, S. (1979) *FEBS* **100,** 313–317.
9. Clementi, E., Corongiu, G., Jonsson & Romano, S. (1980) *J. Chem. Phys.* **72,** 260–263.
10. Lie, G. C., Yoshimine, M. & Clementi, E. (1976) *J. Chem. Phys.* **64,** 2314–2323.
11. Matsuoka, O., Yoshimine, M. & Clementi, E. (1976) *J. Chem. Phys.* **64,** 1351–1361.
12. Clementi, E. (1976) *Lecture Notes in Chemistry,* Vol. 2, Springer-Verlag, Berlin, pp. 1–107.
13. Falk, M., Hartman, K. A. & Lord, R. C. (1962) *J. Am. Chem. Soc.* **84,** 3843–3846.
14. Falk, M., Hartman, K. A. & Lord, R. C. (1963) *J. Am. Chem. Soc.* **85,** 387–391.
15. Falk, M., Hartman, K. A. & Lord, R. C. (1963) *J. Am. Chem. Soc.* **85,** 391–394.
16. Rupprecht, A. & Forslind, B. (1970) *Biochim. Biophys. Acta* **204,** 304–329.
17. Hearst, J. E. & Vinograd, J. (1961) *Proc. Natl. Acad. Sci. USA* **47,** 825–830.
18. Hearst, J. E. & Vinograd, J. (1961) *Proc. Natl. Acad. Sci. USA* **47,** 999–1005.
19. Hearst, J. E. & Vinograd, J. (1961) *Proc. Natl. Acad. Sci. USA* **47,** 1005–1014.
20. Wolf, B. & Hanlon, S. (1975) *Biochemistry* **14,** 1661–1670.
21. Tunis, M. J. B. & Hearst, J. E. (1968) *Biopolymers* **6,** 1325–1344.
22. Tunis, M. J. B. & Hearst, J. E. (1968) *Biopolymers* **6,** 1345–1353.
23. Kuntz, I. E., Branfield, T. S., Law, G. A. & Purcell, G. V. (1969) *Science* **163,** 1329–1331.
24. Privalov, P. L., Ptitsyn, O.B. & Birshtein, T. M.(1969) *Biopolymers* **8,** 559–571.
25. Texter, J. (1978) *Prog. Biophys. Mol. Biol.* **33,** 83–97.
26. Clementi, E., Corongiu, G. & Lelj, F. (1979) *J. Chem. Phys.* **70,** 3726–3729.
27. Corongiu, G. & Clementi, E. (1978) *Gazz. Chim. Ital.* **108,** 687–691.
28. Dahlborg, U. & Rupprecht, A. (1971) *Biopolymers* **10,** 849–863.
29. Clementi, E. (1980) *Computational Aspects for Large Chemical Systems, Lecture Notes in Chemistry, Vol. 19,* Springer-Verlag, Berlin, pp. 1–184.
30. Corongiu, G. & Clementi, E. (unpublished).

Received June 17, 1980
Accepted September 9, 1980

Simulations of the Solvent Structure for Macromolecules. II. Structure of Water Solvating Na^+-B-DNA at 300 K and a Model for Conformational Transitions Induced by Solvent Variations

GIORGINA CORONGIU* and ENRICO CLEMENTI,† *IBM DPPG, Department B28, Building 703-1, P.O. Box 390, Poughkeepsie, New York 12602*

Synopsis

The structure of water and its interaction energy with a fragment of B-DNA composed of 12 base pairs and of the corresponding 24 sugar and 22 phosphate units and Na^+ ions (one at each phosphate group) are analyzed using Monte Carlo simulations. The sample of water molecules, at the simulated temperature of 300 K, is composed of 447 water molecules. The results are discussed either in terms of statistical analyses over the 2,000,000 simulated conformations (after equilibration) or with reference to an "average configuration." Comparison is made to a simulation previously presented for the same system but without counterions. Isotherm at different relative humidity, hydration, and reactivity scales for different sites, the hydration number at each site, the structure of intraphosphate and interphosphate hydrogen-bonded filaments of water are reported and discussed. The stabilization of the B-conformation induced by the solvent with counterion ("ion-induced compression effect") is analyzed on the base of the above findings. A preliminary model to predict conformational transition in DNA is presented. The analyses reported are very detailed to allow refined interpretations of spectroscopic (infrared, Raman, and nmr) and scattering (x-ray and neutron beam) data on DNA in solution.

INTRODUCTION

In the previous paper[1] of this series we have considered aspects of the solvation problem for a B-DNA double-helix fragment composed of 12 base pairs and of the corresponding 24 sugar and 22 phosphate units. Preliminary studies on the solvation of the A- and B-DNA single helix[2,3] have been presented previously. In this work the PO_4^- units have been neutralized with Na^+ counterions. As known, the double helix either in solution or in fibers is stabilized by the presence of counterions. Preliminary data concerning the stabilization induced by Na^+ counterions have been presented in Ref. 1 and in a recent monograph.[4] As previously done, we use the Monte Carlo techniques and *ab initio* derived atom–atom pair potentials, and a

* Permanent address: Istituto G. Donegani, S.p.A., Novara, Italy.
† To whom reprint requests should be addressed.

Biopolymers, Vol. 20, 2427–2483 (1981)

CCC 0006-3525/81/112427-57$05.70

variable number of water molecules (up to 447) to simulate the hydration at several relative humidity values; all computations are performed at a simulated temperature of 300 K.

For the B-DNA fragment's geometrical characterization, we refer to Ref. 1. In this work the Na^+ ion is placed near the free oxygens of PO_4^- at its minimum-energy position (the Na^+ atom–pair-potentials with model compounds containing the PO_4^- group are obtained in our standard way).[4-6] The ion is kept at this *fixed* position during the Monte Carlo simulation; this is a reasonable approximation, because the very strong attraction from PO_4^- and Na^+ yields a well-pronounced, localized, and deep minimum (see last section). Since, however, alternative positions for the Na^+ counterions could be considered, as later discussed in this paper, we present this computation as a *model* study rather than as a *realistic simulation*, which would require the Na^+ counterions positions not to be fixed as input but to be determined by the Monte Carlo technique. Clearly, such simulation would require a notable increase in computer time, not only because 22 particles (the ions) are added to the system, but mainly because the repositioning of one ion brings about the need to relocate its solvation shell. (In other words, one has reason to expect a *very* slow convergence.)

We recall (see Ref. 1) that the water molecules are enclosed in a cylindrical volume, coaxial to the long axis (Z axis) of Na^+-B-DNA double-helix fragment; we also recall that a full B-DNA turn requires 10 base pairs,[7] rather than the 12 we have used (the top and bottom ones are added to render more realistic the boundary conditions in our simulation). The sequence in our fragment is A·T, T·A, G·C, A·T, C·G, T·A, G·C, C·G, A·T, T·A, C·G, G·C (see Ref. 7).

In this paper we discuss the structure of water at high relative humidity (447 water molecules, or about 95% relative humidity[8]), either making use of statistical analyses or with reference to an "average" water configuration. We also discuss the structure of water in the grooves and present an analysis relevant to the adsorption–desorption isotherms and a model for conformation transitions. Lastly, we point out the limitations and approximations of this computational–theoretical treatment.

STRUCTURE OF BOUND WATER AT HIGH RELATIVE HUMIDITY: STATISTICAL ANALYSES

The water structure at high humidity (447 water molecules) is analyzed in greater detail, since it corresponds to a situation expected to be rather near to the one for Na^+-B-DNA in physiological solution; low humidity is of interest mainly for comparison with physical–chemical (rather than biological) experiments, for example, on DNA fibers. The two helix and the constituent chemical groups are differentiated in this paper by the presence or absence of an asterisk; thus, for example, P(n) and P*(n), G(n) and G(n)* refer to the nth phosphate group or to nth base (guanine) of either the h or the h* helix, namely, the two strands forming the double

helix. Water molecules are labeled with an index (from 1 to 447), and generally, a water molecule with an index m has its oxygen atom above (projection on the Z-axis) any water molecule with an index value larger than m. Water molecules are not represented by providing the oxygen and the hydrogen probability density maps (cumbersome for interpretation and unyielding as input data for statistical analysis) but with the more familiar three-sphere representation obtained, however, with the new algorithm explained in Ref. 1. Thus, the ORTEP-like representation does not refer to a single configuration but is *statistically* meaningful, since it is obtained as an average of many thousand configurations; such configuration is hereafter referred to as "*average configuration*." The Monte Carlo *starting* configuration is selected to be the final configuration from the B-DNA study (see Ref. 1), after having carved out a few water molecules in the vicinity of Na^+ (these are initially placed near the boundary of the cylindrical volume). This starting configuration allows us to obtain convergence with about 2000 moves *per water molecule*. The statistical data below reported are obtained by collecting additional 2000–5000 moves per water molecule and by disregarding all the "moves" prior to convergence. For low relative humidity, the attraction by PO_4^- and Na^+ to the relatively few water molecules being very strong, one could expect that the above number of "moves" is unnecessarily too large. This intuitive notion is, however, incorrect, since the sample with few water molecules interacting with the DNA fragments has more freedom of motion (for a given water molecule) than a sample with many water molecules.

In Fig. 1, we present the histogram for the number of oxygen (or hydrogen) atoms, $N(i)$, (i = O or i = H) with a distance R ($i - a$) from the nearest atom a of Na^+-B-DNA. The R ($i - a$) value is obtained by comparing the distances from the water atoms to the Na^+-B-DNA atoms and by selecting the a atom corresponding to the shorter distance. Comparing this histogram, with the equivalent one for B-DNA (see Fig. 1 of Ref. 1), we notice a broad oxygen atom peak with a maximum at about R(O − a) = 2.5 Å and

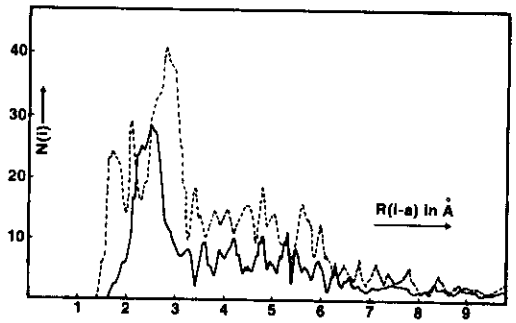


Fig. 1. Distribution of water molecules in Na^+-B-DNA. Distance of oxygen (solid line) and hydrogen atoms (dotted line) from nearest atom in Na^+-B-DNA fragment.

extending up to somewhat more than 3.0 Å and three (no longer symmetrically placed) peaks at about $R(H - a) = 1.8$, 2.1, and 2.8 Å. Taking as first shell all the water molecules up to about $R(i - a) = 3.0$ Å, we conclude that there are more water molecules in Na^+-B-DNAs first shell (with both hydrogen atoms pointing *away* from Na^+-B-DNA) than in B-DNA. The reasons for the hydrogen atom orientation is to be found in the different orientation assumed by a water molecule in the field of Na^+ ions[9-11] relative to the orientation in the field of the PO_4^- group.[5] From Fig. 1 we note a richer fine structure in Na^+-B-DNA than in B-DNA, especially in the region between $R(i - a) = 3.0$ Å and $R(i - a) = 9.0$ Å, suggesting a more structured water pattern; in the region $R(i - a) = 0.0$ Å to $R(i - a) = 3.0$ Å, the integrated area of the distribution $N(i)$ versus $R(i - a)$ is larger than for the equivalent histogram in B-DNA, giving evidences of more densely packed water. The ion-induced *compression* effect is fully expected on the basis of the energetic differences for a water molecule in Na^+-B-DNA, relative to a water molecule in B-DNA as discussed in Ref. 1. Less qualitatively, we obtain for B-DNA,

$$n(O) = \int_{R_1}^{R_2} N(O)\, dR = 176 \qquad \text{for } R_1 = 0.0 \text{ Å and } R_2 = 3.0 \text{ Å}$$

$$n(H) = \int_{R_1}^{R_2} N(H)\, dR = 373 \qquad \text{for } R_1 = 0.0 \text{ Å and } R_2 = 3.0 \text{ Å}$$

and for Na^+-B-DNA,

$$n'(O) = \int_{R_1}^{R_2} N(O)\, dR = 211 \qquad \text{for } R_1 = 0.0 \text{ Å and } R_2 = 3.0 \text{ Å}$$

$$n'(H) = \int_{R_1}^{R_2} N(H)\, dR = 418 \qquad \text{for } R_1 = 0.0 \text{ Å and } R_2 = 3.0 \text{ Å}$$

Clearly, the corresponding integrals from $R_1 = 3.0$ to $R_2 = 14.5$ Å are as follows:

B-DNA

$$m(O) = 447 - n(O) = 271$$

$$m(H) = 894 - n(H) = 521$$

Na^+-B-DNA

$$m'(O) = 447 - n'(O) = 236$$

$$m'(H) = 894 - n'(H) = 476$$

In conclusion, considering a distance up to $R = 3.0$ Å from the nearest atoms of the solute, we find 176 water molecules (and 21 residual hydrogen atoms, belonging to other water molecules) in B-DNA; this number increases to 209 (and two residual oxygen atoms) in Na^+-B-DNA. Alternatively stated, the 22 Na^+ ions in B-DNA have crowded in about 30 additional water molecules into the first solvation shell (or about 1.5 water molecules per

Na^+ ion) relative to B-DNA. It is noted that the distance of 3 Å, obtained from Fig. 1, also appears in the early literature on DNA solvation, particularly in the notable work by Lewin.[12]

An obvious consequence (as we shall later discuss in detail) of the ion-induced *compression* effect is that the double-helix conformation is sterically stabilized relative to a less packed situation (for example, B-DNA without counterions). A second obvious consequence is that the ion-induced compression effect must be some function of the counterion interaction energy with water; thus it will vary from ion to ion. Experimentally, it is well known that temperature, relative humidity, ionic concentration, and specificity are the variables capable of inducing conformational transition in DNA.

A complementary global analysis on the structure of water in Na^+-B-DNA and in B-DNA can be obtained by considering the number of water molecules enclosed in the volume between two coaxial cylinders, with the main axis coincident with the B-DNA fragment's main axis (Z direction). We designate as $R(i)$ and $R(o)$ the radius of the inside and of the outside cylinders, and we select $\Delta R = R(o) - R(i) = 0.2$ Å. The resulting diagrams for the hydrogen probability distribution as function of R (dashed line) and for the oxygen atom (full line) are reported in Fig. 2 for B-DNA and in Fig. 3 for Na^+-B-DNA. These diagrams can also be compared with isoenergy contour maps, previously reported for A-DNA, B-DNA double or single helix[2-4,13-16]

In the figures we report the probability distribution for the 447 water molecules (*"total"* distribution), a partial distribution related to the water molecules bound to *hydrophilic sites* (*"bound"* water molecules distribution) and a second partial distribution (*"remainder"*) defined as the difference between *"total"* minus *"bound"* distributions.

We recall that in the first solvation shell one finds water molecules near to hydrophobic sites, generally acting as bridges between hydrophilic sites. Thus, the *"first solvation shell"* distribution contains the *"bound"* distribution and additional water molecules. The difference between the "total" and "the first solvation shell" distributions provides the *"groove"* distribution. It is clearly a matter of taste to speculate how much a "groove distribution" coincides with "second and third solvation shell distributions." We have opted for the "groove" terminology, since it allows further differentiation between major and minor groove, a distinction of relevance in discussing polynucleic acids. Lewin[12] subdivides the major groove's water molecules into three belts (upper, middle, and lower); our subdivision is about equivalent, since the water molecules of the upper and lower belts correspond essentially to our first solvation shell's water molecules.

In Figs. 2 and 3 we report details of the distribution around specific groups (hereafter referred as *"site analyses"*), namely, the N:, NH_2, CO sites of the bases, the oxygen O1′ of the sugar, the bound oxygen O3′ of PO_4^-, the two free oxygens O2P and O1P of PO_4^-, and Na^+. In the abscissa we report R in Å, and in the ordinate we plot the number of oxygen (or hy-

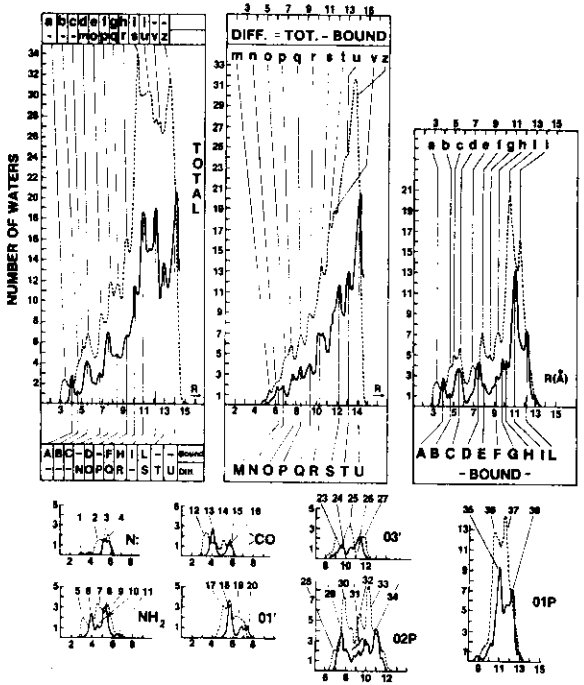


Fig. 2. Probability distributions for (447) water molecule's hydrogen and oxygen atoms as function of R in solvated B-DNA at 300 K. The three inserts at the top refer to total, "bound" water and "remainder" distribution (left to right).

drogen) atoms contained in the previously described volume element defined by $R(i)$ and $R(i) + 0.2$ Å. Since (as previously pointed out, see Refs. 1–6, 14–16) a given water molecule can belong to the first solvation shell of more than one atom (of solute), the distributions of the water molecules for the N:, NH_2, CO, O1′, O3′, O2P, O1P, and Na^+ groups are *not* additive.

The diagrams presented in Figs. 2 and 3 are *"high-resolution"* diagrams obtained by analyzing all the conformations computed in the Monte Carlo simulation; the resolution is about 0.2 Å. *"Low-resolution"* diagrams can be obtained much more easily and cheaply by analyzing the water distribution in the "average conformation" and/or by considering a volume defined by R and $R + 0.5$ Å (as done in Refs. 1 and 4); the resolution is about 0.5 Å. High- and low-resolution analyses are complementary, since the first provides accurate quantitative data of not too easy interpretation, whereas the latter provides less accurate data that can, however, be immediately visualized.

We note that since a water molecule is influenced by the entire field of

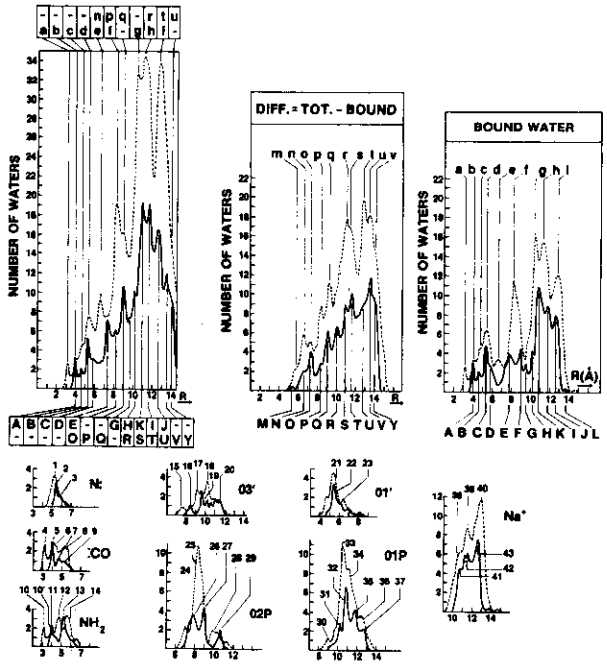


Fig. 3. Probability distributions for (447) water molecule's hydrogen and oxygen atoms as function of R in solvated Na^+-B-DNA at 300 K (see Fig. 2 caption).

the DNA fragment, there is necessarily an element of ambiguity in deciding if a given water molecule "solvates" only a given site, unless an exact definition is provided. Our selection criteria in assigning a water molecule as "bound" to a given atom a of the B-DNA fragment are: (1) the oxygen atom of water and the atom a internuclear separation, $R(a,O)$, must be smaller or equal to a threshold value $T(a,O)$; and (2) one hydrogen atom of a given water molecule and the atom a internuclear separation, $R(a,H)$ must be compared to a threshold value $T(a,H)$ to ensure the proper orientation of that water molecule relative to atom a. In this way we can distinguish the case a—H-O from the case a—O-H [for example N: (in a base) from H of NH_2].

In Fig. 4 we report the projection onto the X-Y plane of the third (A3-T3*) and fourth (C4-G4*) base pairs of our fragment (and the corresponding sugar–phosphate groups) in order to clarify our definitions of bound water, *first* solvation shell water, and groove water. The thick solid line, *composed* by a family of circumferences of radius $T(a,O)$ and *enclosing* either A3-T3* or C4-G4*, represents regions of "bound" water molecules. A number of atoms a have been identified in the figure with a larger dot to represent the nuclear position and the radius origin. Some of the spheres are inside the

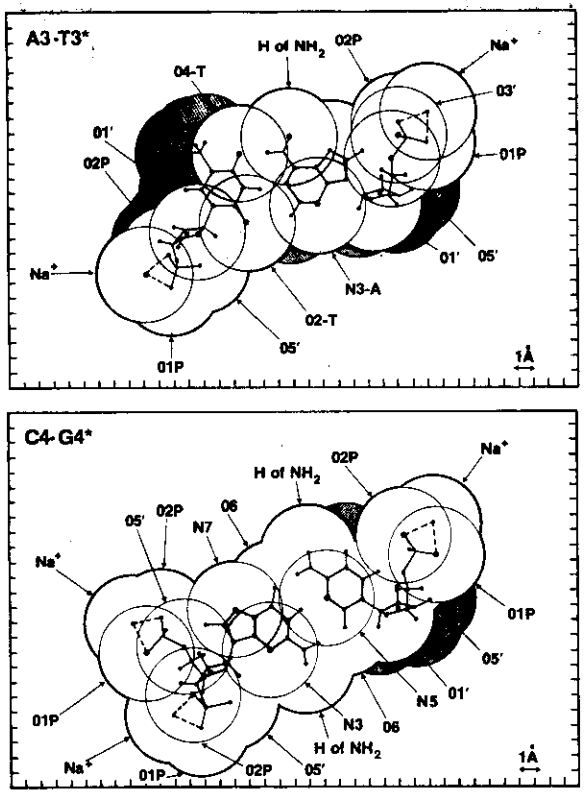


Fig. 4. First solvation shell decomposition. Example for two subunits of the Na^+-B-DNA fragment.

thick solid line, like O3′ in A3·T3* and O2P, O5′, N3, N5 in C4·G4*. One can notice at a glance the strong overlap of these spheres, and this fact constitutes the main reason for the "nonadditivity," namely, the total number of water molecules found to be *bound* to be DNA fragment is smaller than the sum of the numbers of water molecules found to be *bound* to the atoms composing the DNA fragment. However, the "orientation" criterion imposed on the hydrogen atoms [threshold $T(a,H)$] reduces the above inconvenience considerably. The shaded areas at the periphery of A3·T3* and C4·G4* refer to water molecules in the *first solvation shell* but not bound to hydrophilic sites (see, for example, the area for CH_3 of thymine in A3·T3*). Water molecules in this volume might be counted more than once in our analyses, once as being in the first solvation shell (for example, CH_3 for T3*) and once as being bound to a given nearby site. Finally, in C4·G4*, we have presented two PO_4^--Na^+ groups near to guanine, G4*.

Notice the overlap of O1P (at one phosphate group) with Na^+ (on the second group), or the overlap of O2P at one group with O5′ at the second group. Thus, we stress once more that a water molecule *should not be* physically associated to a *single* site; its location and orientation are due to the *entire* field of the fragment. An assignment is, however, essential in interpreting the experimental data, in particular ir, Raman, and scattering data, and also very important in describing and understanding the solvent structure.

In Figs. 2 and 3 the oxygen atom peaks are identified with capital letters, A, B, . . . , U, and the hydrogen atom peaks are identified with lower-case letters a, b, . . . , z. For the site diagrams (nonadditive, see bottom of Fig. 2), the peaks are indicated with numerals.

Before analyzing the distribution data in Figs. 2 and 3, we explain the main features of Figs. 5 and 6. For each probability distribution (Figs. 2 and 3) there is a corresponding energy distribution, reported in Figs. 5 and 6. The energy units are kJ/mol. The minima in the energy distribution are given by capital letters for the "total," "bound," and "remainder" waters, by numerical indices in the "site" analyses. The energy distributions are given as full lines or as dashed lines depending on the statistical significance of the energy. For example, in Fig. 5, the energy distribution for

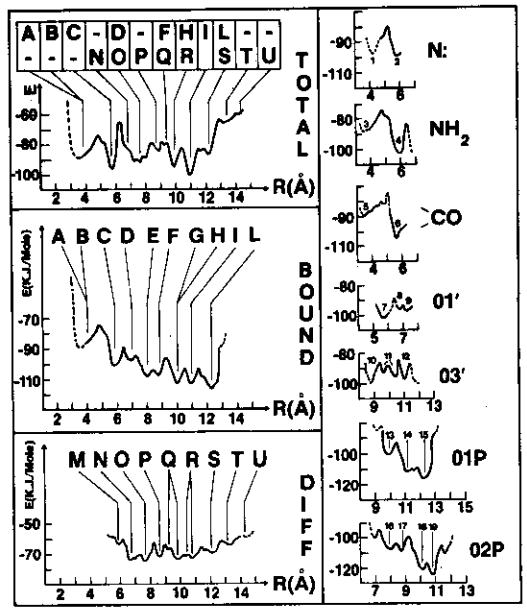


Fig. 5. Energy distributions (in kJ/mol) versus R (in Å) for the water molecules solvating B-DNA at 300 K. Top, middle, and bottom left for "*total*" distribution, "bound" water, and "remainder," respectively; the other inserts refer to specific "site" distributions.

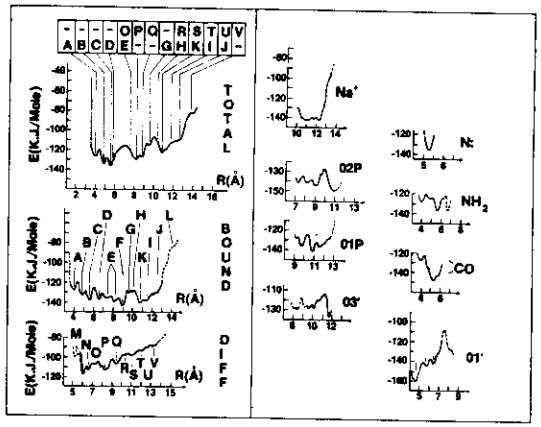


Fig. 6. Energy distribution (in kJ/mol) versus R (in Å) for the water molecules solvating Na^+-B-DNA at 300 K (see caption of Fig. 5).

the N: atoms (bases) shows two minima, designated as 1 and 2; the corresponding probability density distribution (Fig. 2) is low in the interval R = 2–4 Å; therefore the corresponding energy distribution is statistically not too meaningful and is reported as a dotted line rather than a full line.

The analyses of Figs. 2, 3, 5, and 6 are facilitated by recalling the R value of a few important groups or atoms of the DNA fragments (where R is the distance of a given DNA atom from the long axis of DNA). For *adenine* the two hydrogen atoms of NH_2 are at R = 1.8 and 2.8 Å, and the lone-pair nitrogen atoms are at R = 3.2 (N3) and 4.0 Å (N7). For *cytosine* the two hydrogen atoms of NH_2 are at R = 2.1 and 3.7 Å, and the oxygen atom at R = 3.7 Å (O2). For *guanine* the two hydrogen atoms of NH_2 are at R = 3.0 and 4.0 Å, the two lone-pair nitrogen are at R = 3.3 (N3) and R = 3.9 Å (N7), and the oxygen atom at R = 1.7 Å (O6). For *thymine* one oxygen atom is at R = 2.8 Å (O4) and the other is at R = 3.6 Å (O2). These are not the only possible bonding sites for single bases or base pairs in solution. As known, more sites are available in solution[16] for separated base pairs; these are not present in B-DNA, mainly because of steric hindrance (base-pair stacking and the presence of the sugar groups). For the sugar the oxygen atom O1′ is at R = 6.2 Å. For the *phosphate groups* the two free oxygen atoms are at R = 10.2 (O1P) and R = 9.1 Å (O2P), whereas the two bound oxygen atoms are at R = 6.7 (O5′) and at R = 8.8 Å (O3′); we recall that O5′ is near the CH_2 hydrophobic group. The *sodium* ions are placed as previously discussed at R = 10.7 Å.

With all the above in mind, we can now analyze the data of Figs. 2, 3, 5, and 6. Let us start by comparing the total distributions of the 447 water molecules in B-DNA and in Na^+-B-DNA (top-left inserts in Figs. 2 and 3). We notice a set of peaks for the hydrogen and oxygen atoms; the two higher

peaks for the oxygens (peaks K and I) at R = 10.8 and 11.7 Å for Na^+-B-DNA can be compared to the highest peak, U at R = 14.2 Å in B-DNA. The shift confirms the ion-induced *compression effect*, previously mentioned in analyzing Fig. 1 (or forecasted in previous computations, Refs. 1–4, 14). Equally evident is the main overall feature in the hydrogen distribution (relative to the oxygen distribution); the hydrogens are shifted toward smaller R values in B-DNA than in Na^+-B-DNA; in addition, the B-DNA and the Na^+-B-DNA patterns of peaks are different. A peak-by-peak identification cannot be made without reference to the distributions of the remainder water (Figs. 2 and 3, top-middle inserts) and of the bound water molecules (Figs. 2 and 3, top-right inserts). By integration of the "bound" distribution, we find that in B-DNA, 157 ± 2 water molecules out of the 447 are bound water. The remainder water population increases with R (as fully expected), with a characteristic pattern (peaks M to U in B-DNA, M to Y in Na^+-B-DNA). By comparing Na^+-B-DNA with B-DNA we notice that the remainder clearly shows the ion-induced compression effect. This is to be expected; indeed, these molecules can be compressed more readily than the bound molecules, the latter being trapped at the sites, that is, at the most intense and attractive field region, as shown by the energy distributions for the water molecules in the remainder relative to those in the bound (see also Figs. 5 and 6, the three left-hand inserts).

The presence of the counterion brings about another global effect: in Na^+-B-DNA the *bound* water molecules have approximately a constant energy attraction from R = 4.0 to 12.0 Å, whereas in B-DNA the bound water molecules are more attracted at R = 13.0 than at lower values of R. Thus, the counterion brings about an overall increase in the attraction for all the water molecules, and this effect is more prominent for the water molecules near the base pairs than for those at large R values. Another important feature is that the remainder water molecules show the lowest energy at R = 6.0 in Na^+-B-DNA (minima M and N), whereas the lowest energy is between R = 6.0, and 11.0 Å in B-DNA (minima N, P, Q, and R). We draw two conclusions from this observation. *First*, a molecule (for example, a carcinogenic or an anticarcinogenic molecule with polar groups) is expected to be "pressed" toward the base pair *by the field of the counterions*, displacing the groove's water molecules. *Second*, theoretical computations on carcinogenic or anticarcinogenic compounds performed on models where only a single base or a base pair are considered, rather than a full fragment of DNA with its counterions, might not be relevant to problems related to DNA interactions with such compounds, unless, case by case, it has been demonstrated quantitatively that the solvent and counterions effects are small relative to the computed interactions between DNA and the compound (at the temperature considered).

Let us continue with the analyses of the data in Figs. 2–6: in this section we shall focus on the bound water molecules. To understand the density distribution, peak by peak, and the energy minima, we have to consider the sites analyses reported in Figs. 2–6. Let us start with B-DNA (Fig. 2,

top-right insert and the seven small inserts at the bottom). *Peak A* is formed by peak 6 of NH_2 and by peak 13 of CO; *peak B* by peak 11 of NH_2; *peak C* by peak 3 of N:, peak 8 of NH_2, peak 18 of O1′ and peak 16 of CO; *peak D* is formed by peak 29 of O2P and peak 20 of O1′; *peak E* by the right shoulder of peak 29 of O2P (not identified by a number); *peak F* by peak 24 of O3′, peak 31 of O2P (the peaks at R = 8.8, 9.3, 9.8, and 10.2 Å have been labeled collectively as peak 31 in O2P), and the beginning of peak 35 of O1P; *peak G* mainly by peak 31 of O2P, peak 24 of O3′, and peak 35 of O1P; *peak H* by peaks 35 of O1P, 31 of O2P, and 26 of O3′; *peak I* by peaks 35 of O1P (the maximum of this peak), 34 of O2P, and 26 of O3′; and finally *peak L* by peaks 38 of O1P [the low subpeak of L corresponds to the low subpeak of peak 38 of O1P (at R = 13.0 Å)]. Thus, each feature of the "bound" water distribution is now *fully identified*. The same type of analyses can be made for the hydrogen distribution of B-DNA and for the distributions of oxygen and hydrogen atoms in Na^+-B-DNA (supplementary data are available from the authors upon request).

In Tables I–IV we condense the findings so far discussed. In these tables additional site decompositions are presented relative to those given in Figs. 2 and 3, mainly to characterize further those sites known to be of basic importance in problems related to carcinogenic activity and intercalations. We have *excluded* from the analyses the top and the bottom base pairs (AO·TO* and G11·G11*) and the connecting three-phosphate groups and five-sugar units. The exclusion was made in order to delete possibly spurious data due to boundary-condition artifacts. The water molecules at the top and bottom boundaries are designated as *Boundary type 1* or *type 2* depending on whether the binding is of type a-H-O or a-O. Table I presents water molecules bound to specific sites (like O1P or O2P); Table II presents groups of sites (like all the oxygen atoms in PO; Table III presents not only bound water molecules, but all those forming the first hydration shell; and Table IV presents the water molecules around hydrophobic groups containing H atoms (–CH, CH_2, and CH_3) and is given to allow further interpretation of ir or Raman and nmr studies at the carbon atoms. *The information provided by these self-explanatory tables compared to the experimental data accumulated in the last 20 years can be taken as an indication of the evolution in simulation techniques.*

The "sharing" of water molecules at different sites has been pointed out previously in single-helix (A- and B-DNA) solvation studies,[2,3] as well as by Lewin.[12] In this paper, a more definitive answer is provided. For example, the sites at the bases (lines 6–9 in Table II) are solvated cumulatively by 24.56 water molecules, but considering the sites, one by one, 40.09 molecules of water are involved, namely, about 40% (15.53) of the water molecules are "shared." The percentage of sharing in the water bound to the phosphate sites (lines 7–10 and compositive analyses at line 11) is rather small, and we interpret this fact as a consequence of our stringent criteria for the thresholds $T(a - H)$ and $T(a - O)$; relaxing these values we would include "second" solvation shells, where the "sharing" is even more prominent (see Refs. 4 and 11).

TABLE I
Solvation of B-DNA With and Without Counterions at T = 300 K: Water Population Bound to Sites and Its Interaction Energy (kJ/mol)[a]

No.	Site	B-DNA		Na^+-B-DNA		No. of Sites
		Waters	Energy	Waters	Energy	
1	O1P	62.29	−108.18	47.47	−133.54	19
2	O2P	46.44	−111.12	35.91	−141.97	19
3	O5′	1.59	−90.40	0.95	−131.25	19
4	O3′	16.49	−91.09	17.56	−124.92	19
5	Na^+	—	—	65.87	−134.73	19
6	O1′	15.29	−98.68	16.02	−139.87	19
7	NH_2 (A)	5.53	−83.93	4.50	−126.93	5
8	NH_2 (C)	9.84	−79.96	10.56	−120.87	5
9	NH_2 (G)	4.86	−93.80	5.58	−141.66	5
10	N3 (A)	2.27	−85.10	1.78	−136.07	5
11	N7 (A)	0.95	−84.63	1.34	−121.55	5
12	N3 (G)	2.54	−98.23	2.68	−139.84	5
13	N7 (G)	1.16	−83.46	2.45	−120.66	5
14	O2 (T)	1.98	−108.00	2.96	−151.03	5
15	O4 (T)	4.58	−80.10	4.87	−125.50	5
16	O2 (C)	2.00	−98.51	2.00	−142.77	5
17	O6 (G)	5.16	−86.58	5.26	−126.29	5
18	Boundary (type 1)	14.99	−106.94	11.62	−130.72	—
19	Boundary (type 2)	0.00	—	7.78	−130.84	—
20	Total	158.87	−103.87	192.52	−133.75	

[a] In the last column we report the number of times a given site is present in the fragment (excluding the top and bottom, which are considered in the boundaries).

TABLE II
Solvation of B-DNA With and Without Counterions at T = 300 K: Population of Water Bound to Groups of Sites and Its Interaction Energy (kJ/mol)[a]

No.	Groups of Sites	B-DNA		Na^+-B-DNA	
		Waters	Energy	Waters	Energy
1	O1P, O2P, O5′, O3′	113.36	−107.35	93.54	−135.84
2	O1P, O2P, O5′, O3′, Na^+	—	—	141.70	−134.93
3	NH_2 (A, C, G)	17.56	−84.39	18.72	−127.56
4	N: (A, G)	6.92	−89.57	8.25	−130.36
5	O (C, G, T)	11.32	−90.99	12.96	−134.09
6	Adenine	8.45	−83.90	7.62	−128.12
7	Cytosine	11.83	−83.09	12.56	−124.36
8	Guanine	13.25	−91.00	14.81	−132.78
9	Thymine	6.56	−88.52	7.83	−135.15
10	A·T	12.56	−86.60	13.87	−132.55
11	G·C	21.14	−87.15	22.76	−129.18
12	A·T, G·C	24.56	−87.04	30.04	−130.16
13	A·T, G·C, O1′	31.15	−89.20	35.56	−130.17
14	(ph-S) (base-pair) (S-ph)	15.05 + 0.10		19.00 + 0.10	

[a] See Table I (last column) for number of sites (and, therefore, of groups of sites).

TABLE III
Solvation of B-DNA With and Without Counterions at T = 300 K: Population of Water and Its Interaction Energy (kJ/mol) in the First Solvation Shell

No.	"Regions" in First Solvation Shell[a]	B-DNA		Na^+-B-DNA	
		Waters	Energy	Waters	Energy
1	O1P, O2P, O5′, O3′, CH_2	136.78	−104.51	146.10	−133.39
2	O1P, O_2P, O5′, O3′, CH_2, Na^+	—	—	170.09	−133.38
3	O1′	15.36	−98.04	16.21	−139.90
4	Sugar (C_4OH_5)	83.83	−97.79	100.53	−134.21
5	NH_2 (A, C, G)	19.97	−85.93	20.77	−127.85
6	N: (A, G)	7.08	−89.71	8.66	−129.92
7	O (C, G, T)	11.32	−90.99	12.96	−134.09
8	Adenine	11.81	−86.63	13.93	−133.89
9	Cytosine	19.56	−87.33	18.11	−124.48
10	Guanine	19.54	−95.24	20.47	−135.05
11	Thymine	15.63	−87.14	15.33	−132.18
12	A·T	24.37	−87.30	23.93	−132.47
13	G·C	34.84	−91.73	33.40	−130.46
14	A·T, G·C	47.80	−91.14	50.18	−131.08
15	A·T, G·C (C_4OH_5)	106.74	−94.32	122.72	−121.94
16	Total for first solvation shell	185.31	−99.74	223.39	−130.63
17	Total for grooves	261.70	−64.39	223.62	−91.74
18	Total (water–water and water–B-DNA)	447.00	−79.36	447.00	−111.30
	Water–water + total interaction		−15.61 ± 0.05		−22.0 ± 0.05
	Water–B-DNA + total interaction		−63.75 ± 0.10		−89.3 ± 0.1

[a] See last column of Table I.

TABLE IV
Water Molecules Perturbing Hydrogen Atoms (Hydrophobic) in B-DNA With and Without Counterions at T = 300 K[a]

Group	Atoms of Group	B-DNA		Na^+-B-DNA	
		Waters	Energy[a]	Waters	Energy[a]
CH_2	H5'1, H5'2	43.32	−95.12	46.95	−128.30
C_4OH_5	H3, H4	55.97	−95.80	64.17	−131.06
C_4OH_5	H2'1, H2'2	22.13	−99.89	29.69	−137.11
C_4OH_5	H1'	12.44	−98.98	10.80	−139.79
Bases	H in A[b]	3.24	−92.57	4.63	−144.97
Bases	H in C[b]	9.56	−89.08	8.23	−123.30
Bases	H in G[b]	5.00	−104.24	7.09	−134.98
Bases	H in T[b]	10.04	−95.12	9.38	−128.49

[a] kJ/mol.
[b] Hydrogen atoms either involved in the base–base hydrogen bonds or in the NH_2 groups are *not* considered.

By equating *reactivity* of a site, with the reported binding energies of Tables I and II (a rather "extreme" correlation, however, since a study on the interaction of water with DNA, strictly speaking, should discuss only water molecules interacting with DNA), we obtain the following *reactivity* scales:

	B-DNA (300 K)	Na^+-B-DNA (300 K)
NH_2	C < A < G	Same scale
N:	N7(G) ≤ N7(A) ≤ N3(A) < N3(G)	Same scale
O	O4(T) < O6(G) < O2(C) < O2(T)	Same scale
Base	C ≤ A < T ≤ G	C < A < G < T
Pairs	(A·T) ≤ (G·C)	(G·C) < (A·T)

We note, however, that the reactivity between a given molecule and DNA also depends on the sterical possibility for the molecule to interact at one of the above sites. Water molecules are rather small in volume and have a rather large dipole; these two factors must be kept in mind if one uses the above reactivity scales.

In previous papers (Refs. 1–4, 16) we have discussed *hydration scales*; Tables I–III are sufficiently well organized so as to provide such scales

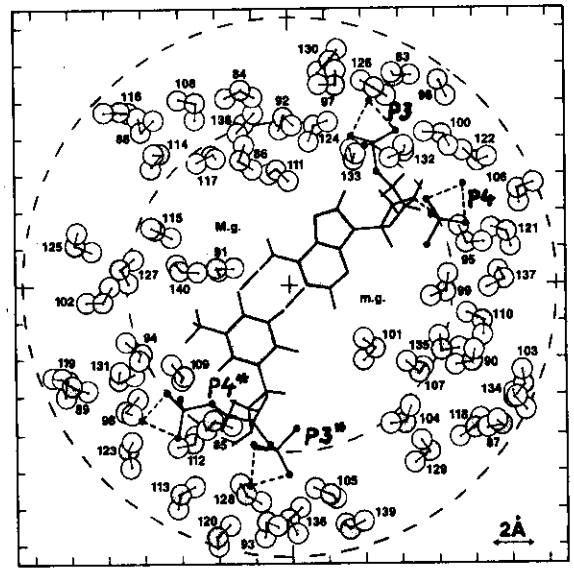


Fig. 7. Water molecules solvating the A3·T3* base pair and its sugar–phosphate–sodium groups; "average" configuration.

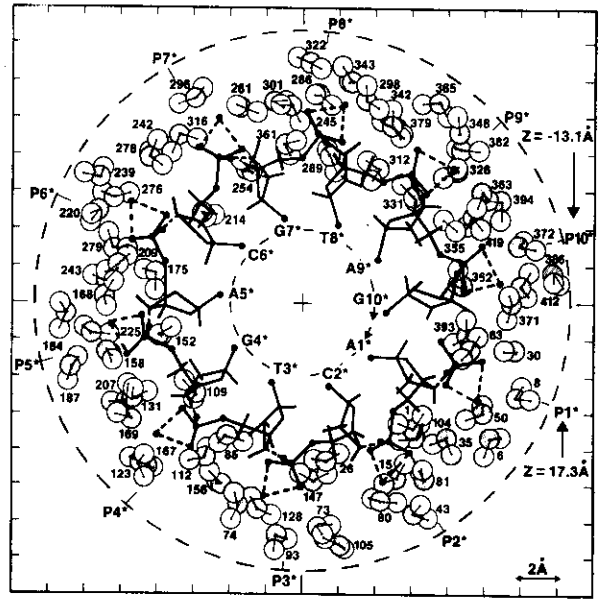


Fig. 8. Water molecules solvating the sodium–phosphate groups of the h* strand of Na^+-B-DNA; "average" configuration.

immediately and to allow comparison between B-DNA and Na^+-B-DNA at 300 K.

STRUCTURE OF WATER AT HIGH RELATIVE HUMIDITY: "AVERAGE CONFIGURATION" REPRESENTATION

Let us now analyze the water structure making use of the "average configuration" data. In Figs. 7 and 8, we analyze the water molecules either included in a disk volume, 4.4 Å thick (from Z = 12.1 to 7.7 Å; see also Fig. 4 in Ref. 1), or the water molecules hydrogen-bonded to the h* helix.

In Fig. 7, we report the third base pair A3·T3* and those water molecules *fully inside the disk*. Several water molecules discussed below are only partly within the disk and are not reported in figure; such molecules, however, are reported in Tables V and VI. To simplify the figure we have omitted the G2·C2* base pair immediately above A3·T3* and the C4·G4* base pair, immediately below. There are four phosphate groups in this disk, P3, P4 (top of figure) and P3*, P4* (bottom of figure; see also Fig. 3). The four Na^+ ions are indicated with a full dot connected to the free oxygen atoms of PO_4^- by dashed lines. The P3 phosphate group is hydrogen-bonded to water molecules 60, 100, 124, 132, and 133; the P4 phosphate group is hydrogen-bonded to water molecules 95, 121, 160, 172, and 185;

the P3* phosphate group is hydrogen-bonded to water molecules 73, 85, 105, 128, and 147; the P4* phosphate group is hydrogen-bonded to the water molecules 109, 112, 131, and 156. The Na^+ ions are solvated as follows: Na^+ (3) is hydrogen-bonded to the water molecules 83, 97, and 126: Na^+ (4) to waters 122, 159, and 160; Na^+ (3)* to waters 74, 93, and 128; and Na^+(4)* to waters 123, 131, 167, and 169. Water molecules 128, 131, and 160 establish a hydrogen bond to both the phosphate and the Na^+ ion. Concerning the base pairs, water molecule 91 is hydrogen-bonded to G2, T3*, and C2*; 111 to A3; and 140 to T3* and G4*. The remaining water molecules of Fig. 7 are either second or third shell solvation and therefore fill the major or the minor groove.

In Fig. 8, we consider the water molecules hydrogen-bonded to the Na^+-PO_4^- groups in the h* helix; in the figure we have indicated the sugar residues and the N atoms of the bases. Comparing this figure with Fig. 3 of Ref. 1, the denser water packing in Na^+-B-DNA relative to B-DNA is apparent.

In Figs. 9–12 we report the water molecules "bound" to the base-pairs or in its immediate neighborhood. One of the base pairs (the one at higher Z) is presented more markedly than the second, which lies immediately below (looking from the Z-axis). From these figures the average position and orientation of the water molecules solvating the base pair can be obtained. These figures are to be compared with Figs. 4–7 of Ref. 1. We recall that the data from the "average configuration" are low-resolution data; on the other hand, these are more immediate than the high-resolution data. The structure accepted by the water molecules represents the effect of the totality of the forces acting on each water molecule, namely, the effect

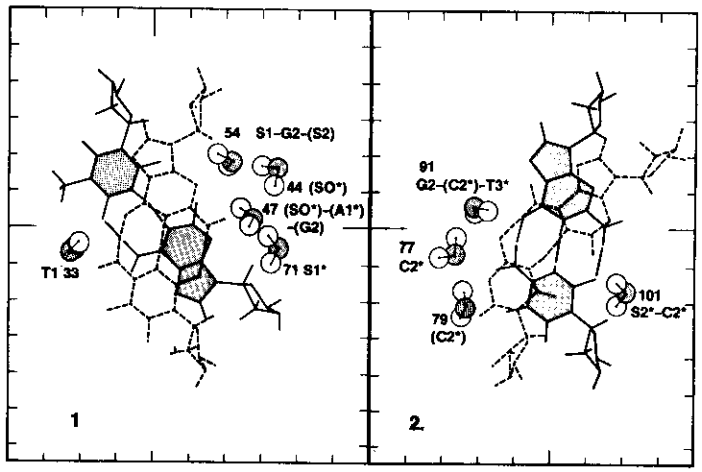


Fig. 9. Hydration of the first three base pairs in Na^+-B-DNA; "average" configuration. (1) T·A* and G·C* pairs; (2) G·C* and A·T* pairs.

TABLE V
Water Molecules Solvating PO_4^- in the h Helix of Na^+-B-DNA[a]

P No.	W No.	Atom (1)	R(H1)	Atom (2)	R(H2)	E(W-Na^+-B-DNA)	E(W-W)
P1	14	O1P	2.1	O1P	3.3	−140.9 ± 5.0	1.0 ± 4.5
	37	O1P	2.0	O1P	3.5	−131.3 ± 5.7	−8.6 ± 3.2
	45	O2P	1.6	O2P	3.0	−119.5 ± 5.7	−13.8 ± 3.2
P2	5	O3′	3.2	O3′	1.8	−106.3 ± 5.0	−6.6 ± 3.8
	29	O1P	2.8	O1P	2.3	−80.1 ± 5.3	−20.0 ± 3.8
	34	O2P	4.0	O2P	3.0	−132.6 ± 2.6	0.6 ± 4.6
	65	O1P	3.0	O1P	1.6	−117.8 ± 6.8	−12.8 ± 4.1
	86	O2P	3.1	O2P	1.8	−123.6 ± 7.6	−22.4 ± 5.1
	92	O1P	2.8	O2P	3.4	−125.4 ± 5.0	−3.3 ± 3.6
P3	60	O3′	2.5	O1P	2.1	−86.0 ± 7.1	−19.8 ± 3.2
	100	O1P	1.6	O1P	3.0	−132.5 ± 5.6	−18.3 ± 4.8
	124	O2P	2.7	O2P	1.7	−128.8 ± 8.8	−17.2 ± 4.3
	132	O1P	2.2	O1P	3.4	−130.9 ± 9.0	−7.2 ± 3.2
	133	O2P	1.9	O2P	3.1	−105.4 ± 3.8	−28.7 ± 4.5
P4	95	O3′	3.2	O3′	2.1	−79.1 ± 8.5	−30.3 ± 2.4
	121	O1P	2.9	O1P	1.7	−111.4 ± 6.5	−24.8 ± 3.5
	160	O2P	2.4	O2P	3.0	−139.3 ± 3.7	−4.7 ± 4.2
	172	O1P	3.1	O1P	1.7	−127.0 ± 11.6	−23.9 ± 3.4
	185	O2P	3.1	O2P	1.8	−114.6 ± 5.0	−28.6 ± 3.4
P5	135	O3′	1.9	O3′	2.9	−90.5 ± 3.6	−30.2 ± 3.1
	196	O1P	1.7	O1P	3.0	−145.2 ± 3.5	−12.4 ± 4.5
	198	O2P	3.0	O2P	1.7	−142.2 ± 3.7	−15.3 ± 4.1
	210	O2P	3.4	O2P	2.7	−139.3 ± 3.7	−5.6 ± 2.8
	233	O2P	1.6	O2P	3.0	−136.9 ± 4.2	−12.4 ± 4.5
P6	190	O3′	3.1	O3′	1.7	−105.1 ± 5.8	−12.1 ± 3.0
	231	O1P	2.8	O1P	1.6	−127.6 ± 4.3	−23.8 ± 2.6
	249	O2P	2.8	O2P	1.7	−140.8 ± 6.8	−12.9 ± 7.4
	250	O1P	3.1	O1P	2.9	−131.2 ± 3.6	−11.0 ± 3.2
	273	O2P	2.7	O2P	1.8	−104.1 ± 4.5	−33.9 ± 3.1
P7	235	O3′	3.1	O3′	1.8	−116.2 ± 5.0	−14.3 ± 3.6
	268	O1P	2.8	O1P	1.5	−132.4 ± 4.2	−6.6 ± 3.2
	287	O1P	2.1	O1P	3.0	−103.6 ± 8.4	−12.0 ± 3.6
	295	O2P	1.7	O2P	3.2	−142.1 ± 6.3	−10.9 ± 3.5
	315	O1P	2.7	O1P	3.2	−85.5 ± 4.8	−31.0 ± 4.1
P8	272	O3′	2.3	O3′	3.6	−119.3 ± 4.1	−17.2 ± 4.5
	297	O1P	1.9	O1P	3.3	−141.6 ± 4.6	−6.4 ± 4.1
	325	O2P	3.0	O2P	1.8	−144.4 ± 5.8	−4.9 ± 3.6
	334	O1P	3.0	O1P	1.6	−139.5 ± 4.0	−11.2 ± 3.5
	351	O2P	1.9	O2P	3.3	−168.4 ± 5.5	−0.5 ± 4.5
P9	347	O1P	1.5	O1P	2.7	−123.8 ± 4.9	−17.5 ± 6.3
	358	O2P	3.4	O2P	2.1	−153.2 ± 5.9	−1.8 ± 4.5
	384	O1P	1.9	O1P	2.8	−119.9 ± 5.5	−15.3 ± 5.0
	389	O2P	3.0	O2P	1.7	−103.1 ± 4.1	−22.8 ± 5.1
P10	360	O3′	2.8	O3′	1.7	−111.5 ± 3.8	−28.8 ± 3.3
	396	O1P	1.8	O1P	2.7	−109.4 ± 4.8	−12.1 ± 2.9
	414	O2P	1.9	O2P	2.9	−113.6 ± 6.1	−13.1 ± 5.0
	420	O1P	1.6	O1P	3.1	−120.9 ± 5.9	−18.1 ± 3.9
	431	O2P	2.5	O2P	2.9	−78.2 ± 6.3	−22.9 ± 3.1
P11	399	O3′	3.2	O3′	1.8	−101.7 ± 3.0	−19.6 ± 3.7
	439	O2P	1.5	O2P	2.9	−138.9 ± 2.6	−15.3 ± 5.1
	442	O1P	1.5	O1P	3.0	−117.9 ± 2.9	−9.4 ± 5.1

[a] The H1 to "atom 1" and H2 to "atom 2" distances are in Å; E in kJ/mol.

TABLE VI
Water Molecules Solvating PO_4^- in the h* Helix of Na^+-B-DNA[a]

P* No.	W No.	Atom (1)	R(H1)	Atom (2)	R(H2)	E(W-Na^+-B-DNA)	E(W-W)
P1*	1	O2P	1.8	O2P	3.0	−118.5 ± 6.4	−7.7 ± 3.2
	30	O1P	2.8	O1P	1.8	−110.4 ± 4.7	−19.4 ± 3.7
	50	O3′	2.9	O3′	3.3	−90.0 ± 4.1	−19.0 ± 3.8
	63	O3′	1.8	O3′	3.1	−109.1 ± 3.2	−18.1 ± 3.7
P2*	26	O2P	3.0	O2P	1.7	−122.7 ± 4.6	−21.5 ± 3.7
	35	O1P	2.9	O1P	1.7	−126.4 ± 5.1	−13.8 ± 5.8
	80	O3′	3.1	O3′	3.4	−134.5 ± 2.5	−8.4 ± 2.8
	81	O1P	2.8	O1P	1.6	−125.8 ± 5.1	−15.5 ± 4.1
	104	O3′	1.8	O3′	3.1	−111.0 ± 3.6	−22.3 ± 4.9
P3*	73	O1P	3.1	O1P	4.1	−72.3 ± 3.0	−31.0 ± 4.6
	85	O2P	2.7	O1P	1.7	−113.1 ± 4.7	−29.3 ± 4.0
	105	O1P	3.0	O1P	1.6	−134.5 ± 5.0	−18.6 ± 4.3
	128	O3′	2.8	O3′	3.1	−128.5 ± 5.0	−14.9 ± 4.9
	147	O3′	1.7	O3′	2.9	−104.8 ± 4.5	−30.9 ± 3.7
P4*	109	O2P	2.8	O2P	1.5	−109.6 ± 6.3	−15.9 ± 5.0
	112	O1P	2.6	O1P	1.9	−112.1 ± 5.6	−13.9 ± 5.4
	131	O2P	2.0	O2P	3.1	−141.3 ± 3.7	−0.0 ± 3.6
	156	O1P	1.7	O1P	3.0	−128.4 ± 5.7	−17.2 ± 3.7
P5*	152	O2P	2.1	O2P	3.4	−130.6 ± 5.0	−6.3 ± 3.4
	158	O1P	3.4	O1P	1.9	−133.7 ± 4.8	−11.8 ± 3.4
	175	O2P	2.4	O2P	1.8	−110.6 ± 4.3	−21.9 ± 3.8
	187	O1P	2.5	O1P	3.4	−84.5 ± 5.3	−25.9 ± 4.1
	207	O1P	3.0	O1P	1.7	−121.2 ± 4.7	−22.9 ± 2.2
	225	O3′	3.5	O3′	2.1	−131.8 ± 8.1	−9.5 ± 6.1
P6*	209	O1P	1.9	O1P	2.9	−113.4 ± 6.8	−25.2 ± 3.5
	214	O2P	1.6	O2P	2.9	−131.3 ± 4.5	−18.0 ± 3.8
	220	O1P	3.5	O1P	2.5	−130.4 ± 5.4	−3.6 ± 3.2
	243	O1P	1.7	O1P	2.9	−128.1 ± 4.3	−7.0 ± 4.0
	261	O2P	2.6	O2P	3.3	−137.8 ± 5.7	−15.5 ± 2.9
	276	O3′	4.1	O3′	3.0	−94.4 ± 6.0	−30.2 ± 3.4
	279	O3′	3.1	O3′	2.1	−96.3 ± 2.5	−22.4 ± 3.3
P7*	254	O2P	1.5	O2P	2.6	−110.2 ± 5.7	−23.1 ± 3.8
	278	O1P	2.9	O1P	1.6	−115.9 ± 6.0	−21.0 ± 3.4
	296	O1P	2.8	O1P	3.6	−130.4 ± 5.1	−15.1 ± 3.6
	316	O3′	3.1	O3′	2.1	−107.5 ± 4.9	−12.6 ± 3.9
P8*	245	O1P	2.7	O1P	1.5	−84.8 ± 2.0	−23.1 ± 2.7
	289	O2P	1.8	O2P	2.7	−134.0 ± 5.3	−20.9 ± 3.1
	301	O1P	1.7	O1P	2.9	−145.4 ± 3.3	−11.5 ± 1.4
	312	O2P	1.8	O2P	2.9	−136.5 ± 5.1	−19.1 ± 5.1
	361	O3′	3.1	O3′	1.9	−115.6 ± 4.1	−19.2 ± 6.1
P9*	331	O2P	2.9	O2P	1.6	−121.8 ± 4.1	−22.3 ± 3.1
	342	O1P	3.0	O1P	1.7	−131.1 ± 6.9	−17.0 ± 4.2
	355	O2P	2.7	O2P	1.7	−120.9 ± 5.3	−18.3 ± 5.1
	365	O1P	2.8	O1P	3.8	−76.5 ± 3.5	−23.0 ± 3.4
	379	O1P	2.8	O1P	1.6	−108.0 ± 4.2	−20.9 ± 3.8
P10*	362	O2P	1.7	O2P	3.1	−120.3 ± 5.7	−13.3 ± 3.8
	393	O2P	1.6	O2P	3.0	−143.7 ± 4.8	−5.9 ± 3.9
	394	O1P	2.7	O1P	2.0	−105.7 ± 6.3	−12.4 ± 3.7
	419	O1P	1.9	O1P	3.2	−103.0 ± 5.1	−16.9 ± 3.1
P11*	403	O2P	2.8	O2P	1.6	−123.9 ± 7.6	−18.6 ± 5.1
	409	O1P	2.6	O1P	1.7	−106.8 ± 14.9	−16.2 ± 3.8
	423	O2P	1.9	O2P	2.5	−125.3 ± 7.8	−5.5 ± 3.5
	446	O1P	1.5	O1P	2.9	−113.0 ± 3.3	−9.0 ± 4.5

[a] The H1 to atom 1 and H2 to atom 2 distances are in Å; E in kJ/mol.

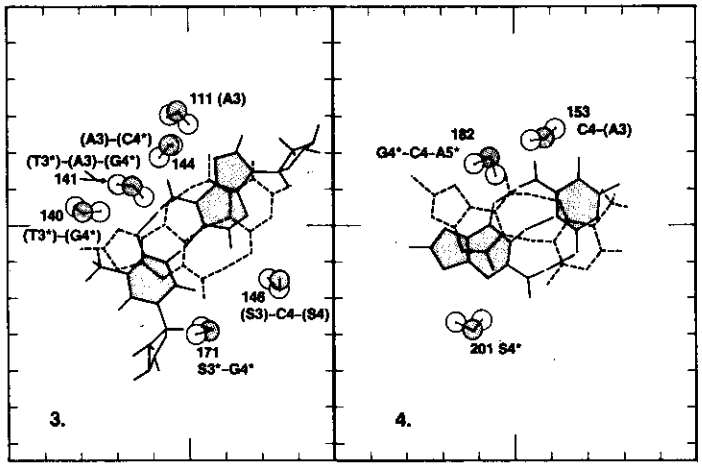


Fig. 10. Hydration of third, fourth, and fifth base pairs in Na^+-B-DNA; "average" configuration. (3) A·T* and C·G* pairs; (4) C·G* and T·A* pairs.

of the interaction energy between a given water molecule and the rest of the solute–solvent system. By shifting the stress from *forces* to *interaction energy* we implicitly restrict ourself to a static representation (Monte Carlo), and in this paper neglect the dynamic representation (molecular dynamics). In the following we provide additional information on the

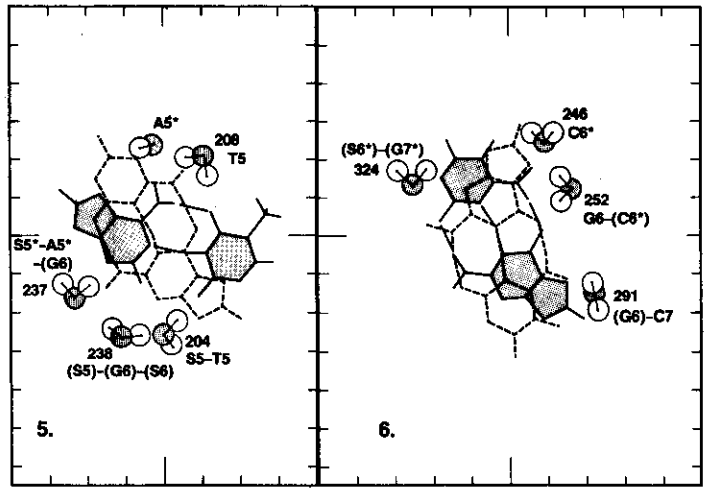


Fig. 11. Hydration in Na^+-B-DNA of base pairs 5, 6, and 7; "average" configuration. (5) T·A* and G·C* pairs; (6) G·C* and C·G* pairs.

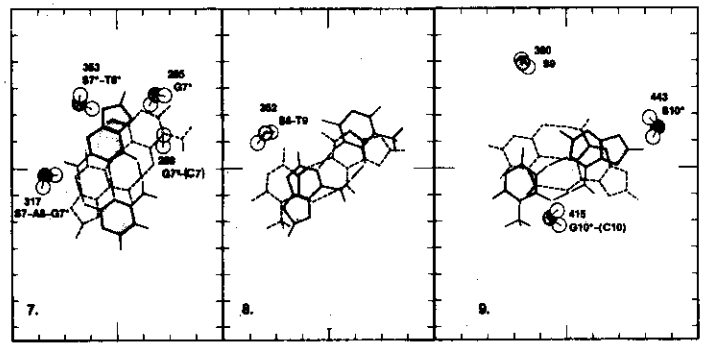


Fig. 12. Hydration in Na^+-B-DNA of base pairs 7, 8, 9, and 10; "average" configuration. (7) C·G* and A·T* pairs; (8) A·T* and T·A* pairs; and (9) T·A* and C·G* pairs.

energy of a given water molecule obtained from the "average configuration," thus complementing the previous energetic analyses reported above. Each "average" energy is obtained by combining a rather wide spectrum of energies, associated to the "average" (and statistically meaningful) position of the water molecule. Since the "average" water molecule is obtained by considering (see Ref. 1) all the water molecules whose hydrogen and oxygen atoms fall within a sphere of 0.5-Å radius centered at either the "average" hydrogen or the "average" oxygen nuclei, the corresponding energy mediates rotations and oscillations around the "average" position; with the selected radius of 0.5 Å, the reported energies correspond to a large sample of water molecules, and therefore, the average standard deviation is large.

This analysis should illustrate how easily one could be deceived by performing computations, however accurate, but with neglect of temperature averaging. Indeed, the energy differences associated with the "average" water molecule (namely, the average standard deviation between the many waters represented by the "average" water) can be of the same magnitude as energy differences often assumed to be biologically significant.

In Tables V–VIII we provide data for the interaction energy of the first shell water molecules solvating the PO_4^- groups in the h (Table V), and in the h* helix (Table VI), the Na^+ ions in the h (Table VII) and h* helixes (Table VIII); we then add the analyses of water molecules for the sugar unit (Table IX) and the base pairs (Table X). In the tables we follow the following format; first, we give the site examined; then an index for the water molecule, consistent with the figures index; then geometrical specification on the hydrogen bond; then the water–total system interaction energy and its mean standard deviation; and finally, the water–water interaction energy and its mean standard deviation. For example, considering Table V, since we deal with the interaction between a phosphate group and water molecules, the geometrical data of interest are the distances $R(H1)$ and $R(H2)$ between the water's hydrogen atoms and the nearest oxygen atom on the

TABLE VII
Water Molecules Solvating the Na^+ Counterions in the h Helix of Na^+-B-DNA[a]

Na^+ (P No.)	W No.	R(O–NA)	Bridges	E(W–Na^+-B-DNA)	E(W-W)
Na^+ (P1)	14	2.1	P1	−140.9 ± 5.0	1.0 ± 4.5
	17	2.0		−121.8 ± 5.9	−13.8 ± 3.3
Na^+ (P2)	34	2.1	P2	−131.6 ± 2.6	0.6 ± 2.6
	55	2.3		−114.3 ± 7.7	−8.7 ± 4.5
	84	2.1		−137.1 ± 3.7	−13.3 ± 3.6
Na^+ (P3)	83	2.3		−129.4 ± 3.5	−18.0 ± 5.5
	97	2.1		−125.3 ± 5.0	−13.8 ± 3.6
	126	2.2		−128.5 ± 3.7	−15.9 ± 4.8
Na^+ (P4)	122	2.2		−135.9 ± 2.9	−9.4 ± 2.4
	159	2.2		−105.4 ± 3.7	−14.8 ± 3.5
	160	2.2	P4	−139.3 ± 3.7	−4.7 ± 4.2
Na^+ (P5)	180	2.3		−110.2 ± 6.7	−14.8 ± 4.9
	210	2.1	P5	−139.3 ± 3.7	−5.6 ± 2.8
	224	2.4		−128.2 ± 3.6	−18.6 ± 3.3
Na^+ (P6)	232	2.2		−130.7 ± 3.5	−21.5 ± 4.0
	250	2.4	P6	−131.2 ± 3.6	−11.0 ± 3.2
	266	3.2		−65.8 ± 5.8	−33.8 ± 4.8
Na^+ (P7)	265	2.2		−143.2 ± 4.7	−10.5 ± 4.2
	270	2.3		−128.8 ± 5.3	−11.3 ± 2.3
	311	2.4		−132.2 ± 5.7	−16.7 ± 3.8
	315	3.1	P7	−85.5 ± 4.8	−31.0 ± 4.1
Na^+ (P8)	302	2.2		−144.9 ± 5.6	−14.8 ± 3.6
	328	2.2		−126.2 ± 6.4	−6.6 ± 3.6
	297	3.1	P8	−141.6 ± 4.6	−6.4 ± 4.1
	249	3.1		−140.8 ± 6.8	−12.9 ± 7.4
	351	2.5	P8	−168.4 ± 5.5	−0.5 ± 4.5
Na^+ (P9)	339	2.4		−134.4 ± 2.5	−10.4 ± 3.7
	358	2.1	P9	−153.2 ± 5.9	−1.8 ± 4.5
	381	2.4		−123.9 ± 6.0	−13.4 ± 4.1
Na^+ (P10)	376	2.2		−109.5 ± 7.6	−14.9 ± 4.7
	395	2.7		−111.3 ± 7.7	−16.3 ± 2.7
	413	2.0		−123.9 ± 3.5	−19.5 ± 3.8
	420	2.9	P10	−120.9 ± 5.9	−18.1 ± 3.9
Na^+ (P11)	410	2.4		−129.0 ± 2.6	−16.1 ± 3.1
	430	2.3		−101.1 ± 9.8	−5.5 ± 2.9
	444	2.3		−117.2 ± 3.2	3.9 ± 4.5

[a] R, in Å; E, in kJ/mol.

TABLE VIII
Water Molecules Solvating the Na^+ Counterions in the h* Helix of Na^+-B-DNA[a]

Na^+ (P* No.)	W No.	R(O–NA)	Bridges	E(W–Na^+-B-DNA)	E(W-W)
Na^+ (P1*)	6	2.3		−112.1 ± 4.9	−9.6 ± 5.1
	8	2.3		−123.1 ± 4.6	−3.7 ± 3.9
Na^+ (P2*)	15	2.4		−109.9 ± 7.4	−20.6 ± 3.0
	43	2.4		−110.5 ± 11.9	−8.8 ± 5.3
	80	2.2	P2*	−134.5 ± 2.5	−8.4 ± 2.8
	81	3.1	P2*	−125.8 ± 5.1	−15.5 ± 4.1
Na^+ (P3*)	74	2.2		−130.1 ± 5.2	−14.9 ± 4.9
	93	2.3		−122.5 ± 5.0	0.5 ± 3.5
	128	2.5	P3*	−128.5 ± 5.0	−14.9 ± 4.9
Na^+ (P4*)	123	2.2		−126.1 ± 3.2	−11.8 ± 3.8
	131	2.4	P4*	−151.3 ± 3.7	0.0 ± 3.6
	167	2.5		−104.9 ± 8.5	−7.1 ± 4.1
	169	2.4		−123.6 ± 7.3	−10.8 ± 3.6
Na^+ (P5*)	158	2.8	P5*	−133.7 ± 4.8	−11.8 ± 3.4
	168	2.2		−132.7 ± 6.3	−1.1 ± 3.8
	184	2.3		−115.0 ± 7.6	−20.7 ± 4.5
	187	3.1	P5*	−84.5 ± 5.3	−25.9 ± 4.1
	225	2.5	P5*	−131.8 ± 8.1	−9.5 ± 6.1
Na^+ (P6*)	220	2.3	P6*	−130.4 ± 5.4	−3.6 ± 3.2
	239	2.1		−124.2 ± 5.8	−7.7 ± 3.3
	242	2.8		−132.6 ± 6.8	−10.2 ± 3.6
	276	3.2	P6*	−94.4 ± 6.0	−30.2 ± 3.4
Na^+ (P7*)	261	2.1	P6*	−137.8 ± 5.7	−15.5 ± 2.9
	296	2.2	P7*	−130.4 ± 5.1	−15.1 ± 3.6
Na^+ (P8*)	286	3.2		−75.7 ± 5.9	−32.1 ± 3.2
	298	2.2		−125.4 ± 4.9	−16.9 ± 3.4
	343	2.2		−126.1 ± 6.1	−15.4 ± 3.2
	322	3.0		−102.0 ± 7.1	−14.4 ± 3.2
	342	3.2	P9*	−131.1 ± 6.9	−17.0 ± 4.2
	326	3.2		−107.3 ± 4.0	−25.3 ± 4.5
Na^+ (P9*)	348	2.2		−127.2 ± 4.4	−12.0 ± 3.4
	363	2.3		−140.0 ± 3.6	−10.7 ± 3.5
	382	2.5		−118.2 ± 9.5	−13.3 ± 3.6
Na^+ (P10*)	371	2.3		−131.4 ± 2.6	−16.9 ± 3.3
	372	2.9		−76.9 ± 6.1	−29.9 ± 3.8
	386	3.1		−68.0 ± 6.0	−10.7 ± 4.0
	412	2.1		−127.3 ± 3.8	−18.3 ± 3.9
Na^+ (P11*)	409	2.9	P11*	−106.8 ± 14.9	−16.2 ± 3.8
	411	2.2		−132.6 ± 4.4	−11.3 ± 4.5
	428	2.4		−108.9 ± 8.5	−15.6 ± 4.5
	441	2.1		−119.7 ± 3.4	−8.3 ± 3.1

[a] R, in Å; E, in kJ/mol.

phosphate. Looking at the first row of Table V, we learn that the phosphate group P1 is hydrogen-bonded to water 14, which has one of the hydrogen atoms at 2.1 Å from O1P and the second hydrogen atom at 3.3 Å from O1P; these are the shortest distances between the hydrogen atoms on water and an oxygen of PO_4^-. The water 14 is attracted to the Na^+-B-DNA fragment by an interaction of 140.9 ± 5.0 kJ/mol, and its overall interaction with the remaining 446 water molecules adds up to a repulsion of 1.0 ± 4.5 kJ/mol, yielding a total attraction of 139.9 ± 9.5 kJ/mol. The standard deviation from the mean value is obtained by considering the many water molecules that are used to define a water molecule in the average conformation. In Tables VII and VIII the relevant geometrical characterization is the distance between Na^+ and the oxygen atom of water; when the water molecule is not only in the first solvation shell of Na^+ but is also hydrogen-bonding a phosphate group, then an annotation is given in the fourth column.

For the sugar units (Table IX), often the same water molecule is hydrogen-bonded to the sugar [S(n) or S(n)*] and at the same time to a base, as found from the previous "statistical" analyses. For example, in the first row of Table IX, water 54 is hydrogen-bonded to the oxygen atom of the sugar residue S1 (2.0 Å), to the sugar residue S2 (2.5 Å), and to the base G2 (2.1 Å). Water 54 is designated as S1-G2-(S2) to evidence the fact that the distance of water 54 to S2 is rather large, and thus the interaction is rather small. This notation has been explained in detail in Ref. 1, where the equivalent to Tables V–X is given for B-DNA (without counterions).

For the base pairs (Table X) we find again that, in general, a water molecule is hydrogen-bonded to more than a single base. For example, in the third row of Table X, water 91 is hydrogen-bonded to G2, T3*, and C2*; the distances between the three atoms of water and the nearest atom in G2, T3*, and C2* are 1.8, 2.2, and 2.3 Å, respectively. In Figs. 9–12 we provide a graphical representation for the water molecules interacting with the base pairs (and with the sugar units). Each insert contains two successive base pairs, the first one with full lines to indicate bonds and dashed areas for the rings, the second one (located below the first one) with dashed lines to indicate bonds.

The overall energetic characterization for the water molecules solvating Na^+-B-DNA obtained from the "average conformation" is reported in Fig. 13 (to be compared with Fig. 9 of Ref. 1), where the interaction energy is reported for the water molecules enclosed in the interface of two coaxial cylinders differing by 0.5 Å in the R radius. Approximately, five energy minima are noted at R values 3.5, 5.0, 7.5, 9.0, and 10.5 Å; two shallow minima are at R = 11.5 and 12.5 Å. As done in Ref. 1, the first three are assigned to the interaction of water with the base pairs, base pair and sugar bridges, and sugar units. The minima at 9.0 and 10.5 Å should be compared with the single minimum found for B-DNA at 11.0 Å; the two shallow minima should be compared with the clearly defined minimum at 12.0 Å in B-DNA. The region R = 8.5 to 12.5 Å includes not only the water mol-

TABLE IX
Water Molecules Solvating the Sugar Units in Na^+-B-DNA[a]

W No.	Atom (1)	R	Atom (2)	R	Atom (3)	R	E(W-Na^+-B-DNA)[a]	E(W-W)	Name
54	S1	2.0	G2	2.1	S2	2.5	−120.0 ± 5.5	−15.3 ± 4.6	S1-G2-(S2)
71	S1*	1.8					−115.9 ± 2.4	−21.5 ± 4.2	S1*
101	S2*	2.2	C2*	1.9			−132.9 ± 3.3	−16.8 ± 5.0	S2*-C2*
171	S3*	1.8	G4*	1.9			−129.2 ± 3.4	−22.9 ± 5.3	S3*-G4*
201	S4*	1.6					−134.7 ± 3.4	−10.9 ± 2.7	S4*
204	S5	1.9	T5	1.9			−141.8 ± 4.2	−21.8 ± 6.3	S5-T5
237	S5*	2.1	A5*	2.1	G6	2.3	−144.7 ± 3.4	−7.6 ± 4.5	S5*-A5*-G6
238	S6	2.4					−124.1 ± 5.1	−8.0 ± 4.5	S6
317	S7	1.8	A8	2.2	G7*	2.1	−132.9 ± 4.9	−10.9 ± 3.8	S7-A8-G7*
352	S8	1.6	T9	2.1			−121.2 ± 4.6	−22.6 ± 3.8	S8-T9
353	S7*	1.7	T8*	2.1			−131.9 ± 4.4	−14.6 ± 4.5	S7*-T8*
380	S9	1.7					−115.0 ± 4.0	−15.4 ± 4.0	S9
443	S10*	1.7					−100.8 ± 3.1	−3.1 ± 4.5	S10*

[a] R, in Å; E, in kJ/mol.

TABLE X
Water Molecules Solvating Bases or Base Pairs in Na^+-B-DNA[a]

W No.	Atom (1)	R	Atom (2)	R	Atom (3)	R	E(W-Na⁺-B-DNA)[a]	E(W-W)	Name
33	T1	1.8					−101.7 ± 2.0	−22.0 ± 2.4	T1
77	C2*	2.1					−79.0 ± 3.5	−30.0 ± 3.6	C2*
91	G2	1.8	T3*	2.2	C2*	2.3	−125.1 ± 4.2	−5.6 ± 2.6	G2-T3*-(C2*)
146	C4	1.8	S3	2.6	S4	2.8	−130.3 ± 3.8	−9.5 ± 2.6	(S3)-C4-(S4)
153	C4	1.7	A3	2.4			−103.2 ± 2.2	−22.8 ± 3.5	C4-(A3)
182	G4*	2.1	C4	2.9	A5*	2.8	−122.7 ± 2.5	−17.9 ± 3.8	G4*-C4-A5*
208	T5	1.9					−119.2 ± 1.7	−13.1 ± 3.8	T5
222	A5*	1.9					−102.5 ± 3.5	−24.0 ± 50	A5*
246	C6*	2.2					−70.6 ± 2.7	−37.4 ± 3.5	C6*
252	G6	2.0	T5	2.8	C6*	2.6	−104.3 ± 5.3	−26.2 ± 2.9	G6-(C6*)-(T5)
285	G7*	2.2					−81.4 ± 4.0	−37.9 ± 3.5	G7*
288	G7*	2.1	C7	2.9			−101.5 ± 5.3	−21.7 ± 4.5	G7*-(C7)
291	C7	1.8	G6	2.3			−112.5 ± 3.4	−20.0 ± 4.5	C7-(G6)
415	G10*	1.9	C10	2.6			−101.3 ± 3.4	−6.9 ± 4.5	G10*-(C10)
47	A1*	2.3	G2	2.6	SO*	2.3	−105.8 ± 2.0	−21.6 ± 3.8	(SO*)-(A1*)-(G2)
79	C2*	2.6					−81.2 ± 18	−31.6 ± 4.1	(C2*)
111	A3	2.4					−99.4 ± 2.5	−22.5 ± 3.4	(A3)
140	T3*	2.8	G4*	2.7			−99.6 ± 4.5	−18.8 ± 3.9	(T3*)-(G4*)
141	T3*	2.5	A3	2.7	G4*	2.3	−104.7 ± 4.4	−18.5 ± 5.0	(T3*)-(A3)-(G4*)
144	A3	2.7	C4	2.9			−86.3 ± 4.6	−30.6 ± 5.6	(A3)-(C4)
238	G6	2.4	S5	2.7	S6	2.4	−124.1 ± 5.1	−8.0 ± 4.5	(S5)-(S6)-(G6)
324	G7*	2.5	S6*	2.8			−123.1 ± 3.7	−22.5 ± 4.5	(S6*)-(G7)

[a] R, in Å; E, in kJ/mol.

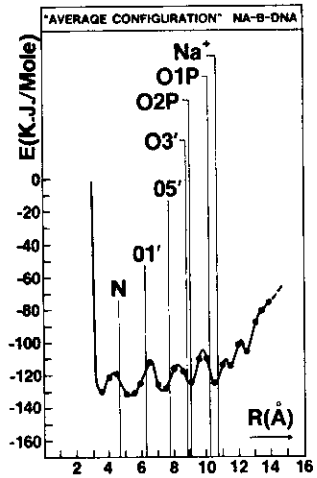


Fig. 13. Interaction energy for the 447-water-molecule simulation; "average" configuration.

ecules around the Na^+-PO_4^- groups, but also the water molecules in the major and minor groove regions.

Comparing Figs. 5 and 13, we note a marked loss of information in the latter. Figure 13, obtained as an average from the many thousand Monte Carlo conformation, represents the "best possible single conformation" (with temperature effects included). Today's literature on "theoretical" studies for solvation of biomolecules often neglects temperature averaging and makes use of a "single configuration" (see, for example, Refs. 17, 18); thus, the disagreement between such studies and the type of data given in Fig. 5 is even more pronounced than the differences between the data of Fig. 5 and those of Fig. 13. In addition, the literature often considers *too small* a fragment and *too few* water molecules. For example, Pullman and his group[19] have analyzed the solvation of a three base-pair fragment of B-DNA. The water molecules were placed *one after the other* at one-half of the fragment and *not all at once*. The implicit assumption of these authors is that the position for one water molecule does not depend on the positions of the following ones, but mainly depends on the positions of the previously placed water molecules. As known (see Ref. 20), this assumption is incorrect. Since available experimental data did report that 4–6 water molecules solvate a phosphate group, 17 water molecules were selected[19]; five were placed at each phosphate group and the remaining two were placed at the three base pairs. Unfortunately, this starting configuration biases the final geometry output. Therefore, this simulation reports mainly "boundary effects" (since the fragment is too small) at low-intermediate relative humidity (since 17 water molecules are too few to describe high humidity) and with temperature neglected. This type of computational

study yields structural data not dissimilar from those obtained using "sticks-and-ball" models, as clearly demonstrated in the early work by Lewin.[12]

STRUCTURE OF WATER IN THE GROOVES

We recall that in Ref. 1 we have detected *intraphosphate* water filaments [connecting P(n) to P($n + 1$) or P*(n) to P*($n + 1$)] and *interphosphate* water filaments [connecting P(n) to P*(m)]. In Na^+-B-DNA the ion-induced compression effect brings about an added feature relative to B-DNA, namely, we observe hydrogen-bond crosslinking between interphosphate water filaments.

In Fig. 14 we examine the structure of water in the minor and major grooves, selecting waters with $8 \leq R \leq 13$ Å. The water molecules are

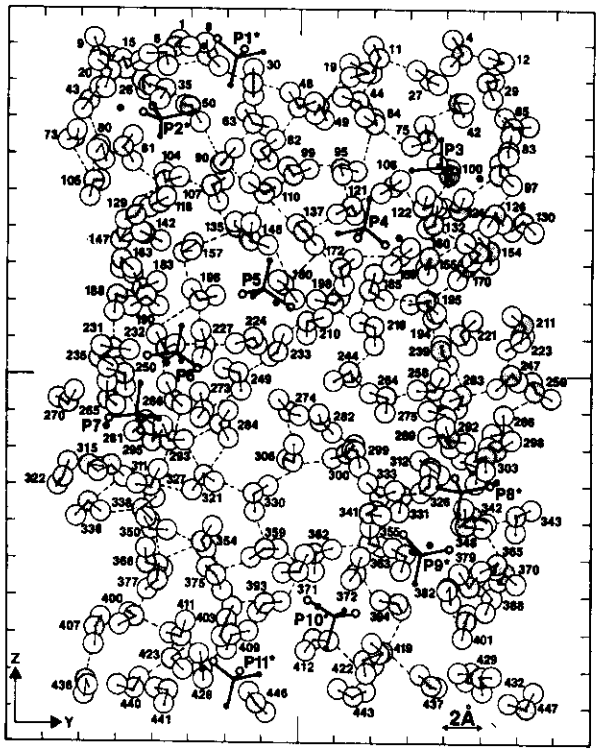


Fig. 14. Major and minor groove waters for the "average" configuration (from R = 8.0 to 13.0).

projected onto the Y-Z plane, namely, a plane containing the long axis (Z-axis) of DNA (see Fig. 15). We have selected those water molecules with positive value for the X coordinate of the oxygen atom. The figure is seen from the $+X$ direction; the five atoms forming a phosphate group are reported; full dots represent the two bound oxygen atoms O3′ and O5′, whereas open dots represent the two free oxygen atoms O1P and O2P. The Na^+ ion is represented by a full dot. The space enclosed by two lines passing through P3, P4, P5, P6, and P7 and through P8*, P9*, P10*, P11* corresponds to the major groove; above and below it are portions of the minor groove. Dotted lines between atoms of water molecules indicate a water–water hydrogen bond.

For Na^+-B-DNA the solvation pattern can be described either in terms of *crosslinked* interphosphate water filaments or in terms of local *cyclic structures* composed of about 5 ± 2 water molecules. For example, the water molecules 293, 321, 330, 359, 371 can be considered as an interphosphate filament; a second filament is formed by the water molecules 293, 327, 338, 350, 354, 375, 403. The crosslinking water molecules for these two interphosphate filaments are water 354 and 375. Alternatively, cyclic structure representations are given, for example, by the water molecules 327, 321, 330, 359, 375, 354, 350, 338 or 375, 359, 371, 393, 403. At this stage we do not see any reason to prefer the crosslinked filament "model" over the cyclic structure "model," and therefore, we rather stress the equivalence of the two alternative descriptions. These findings and those previously reported (Ref. 4, pp. 152–175) can be compared with the data by Neidle et al.,[21] keeping in mind, however, that our base-pair sequence is different and our DNA fragment is not perturbed by intercalated molecules.

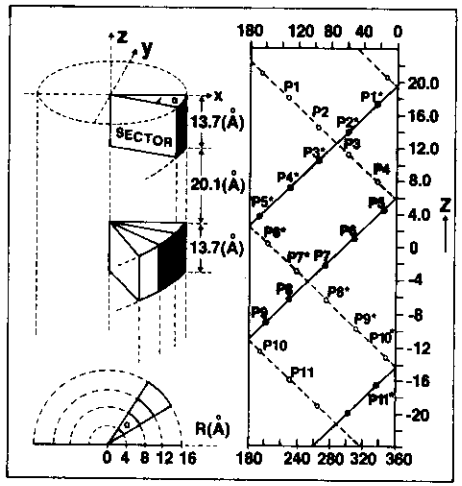


Fig. 15. Grooves analyses: sectors and subvolumes (left), major and minor (dashed) grooves (right); the bottom scale is seen from the $-Y$ axis, the upper scale from the Y axis.

The "low-resolution" structure of water molecules in the grooves is now complemented with the "high-resolution" obtained not from the "average" configuration, but from a full statistical analysis on the 2×10^6 Monte Carlo configurations.

First, we partition the water molecules into those belonging either to the major or minor groove. The cylinder containing the solvent–solute system is subdivided into sectors (see Fig. 15, left insert). For a full B-DNA turn there are N sectors of angle α and a height (along Z) equal to 13.72 Å for the minor groove and 20.08 Å for the major groove; for a full B-DNA turn there are $360/\alpha = N$ sectors (in Fig. 15 we have represented five sectors in the minor groove). We select a clockwise rotation for α (with $\alpha = 0$ for $y = 0$ and x positive, see left side of Fig. 15). Finally, to increase the resolution in locating a water molecule in the grooves, each sector is partitioned into four subvolumes, as indicated at the bottom left insert of Fig. 15, by subdividing R into four segments, namely $0.0 \leq R < 4.0$ Å, $4.0 \leq R < 8.0$ Å, $8.0 \leq R < 12.0$ Å, and $12.0 \leq R < R$ (max) = 14.5 Å. On the right insert of Fig. 15 we report the phosphorous atoms of the h and h* strands; these atoms are on a cylindrical surface and have been projected onto the XZ plane. When α increases from 0° to 180° (or 180° to 360°), we consider all those water molecules with a negative (or positive) value for Y. We count the water molecules which fall into the four subvolumes of a given sector, neglecting, however, those assigned to the first solvation shell. (Therefore, only true "groove" molecules are considered.) After several trials we have selected $\alpha = 4°$ as the value yielding a clear graphic representation, which is reported for the major groove in Fig. 16. [The values reported on the abscissa follow the definition used in Fig. 15; the dashed vertical lines identify the position for the P* atoms (of the h* strand), and the markings at the top of the figure identify the position for the P atoms

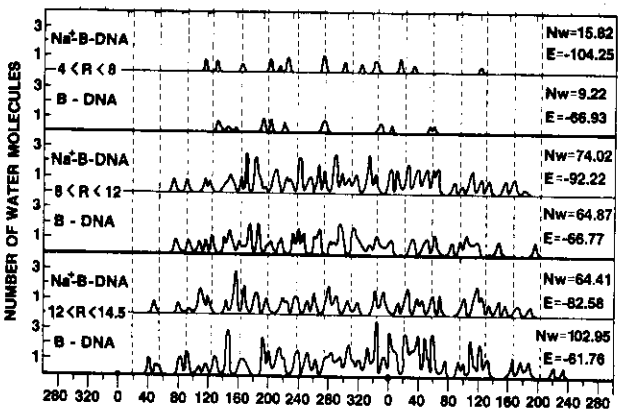


Fig. 16. Probability distributions for water molecules in B-DNA and Na^+-B-DNA in the major groove (see Fig. 15): only groove water molecules are considered.

(of the h strand).] In the figure we compare Na^+-B-DNA with B-DNA. In the region $0.0 \leq R < 4.0$ Å there is no groove water, as previously pointed out, for example, by considering isoenergy maps (see Refs. 4, 15, 16).

The number of water molecules (N_w) and the average interaction energy per water molecule, E (in kJ/mol) varies with R reported in Table XI, from which data we learn a new feature of the packing of the water molecules in the grooves. The interaction energy in the minor groove is larger than in the major groove. (This is expected, because the attractive sites are closer in the minor groove.) For B-DNA the interaction energy is nearly constant in the major groove, as previously predicted,[14,16] but it increases sharply by decreasing R for Na^+-B-DNA. Thus, the groove can be seen as channels to convey reactants to the base pairs; the "flow" in such channels depends both on the counterion position and specificity and on sterical factors of the reactant.

In Fig. 16 a more detailed representation of the "groove" water molecules in the major groove is given. An equivalent diagram for the minor groove shows the same periodicity of peaks but with lower intensity; these peaks represent the "intraphosphate" filaments previously discussed.[1] The gross characterization of the intensity distribution given in Fig. 16 is the nearly periodic existence of peaks, approximately evenly spaced ($\Delta\alpha = 10°$), well developed in the regions $8 \leq R < 12$ and $12 \leq R < R$ (max) and both weakly and inconsistently developed in the region $4 \leq R < 8$ Å. The pattern of peaks clearly points out the existence of a structural organization which, even if temperature-dependent, is not destroyed by thermal motion at room temperature. Therefore, this pattern provides a strong support for the old proposals by Szent Gyorgyi concerning an "ice"-type structure for waters solvating biomolecules. The intensity of the peaks (i.e., the number of the water molecules bound to the phosphates) adds up to "filaments" of about 5–7 water molecules. Since the filament is not fully oriented along the Z-axis (see Fig. 14), a single peak should not be associated to one filament, but it is likely that two to three contiguous peaks participate to the build-up of one "filament." Again we recall the semantic question related to the use of terms "filament" and "cyclic structure" pointed out at the beginning of this section.

The number of water molecules, N_w, and the average interaction energy per water molecule, E, considering not only the groove water but also the first solvation shell water molecules, when partitioned into the sector's subvolumes, are as reported in Table XII. There one notices that the distinction between major and minor groove extends outside the dimension of the solute. Remembering that the Na^+ ions are at the periphery of B-DNA (at $R = 10.7$ Å), the groove structure extends into the liquid surrounding DNA. We shall call "classical groove volume" the one up to $R = 12.0$ Å and "extended groove volume" the one starting at $R = 12.0$ Å. The latter ends when $E = -36$ kJ/mol, namely, the bulk water value with R equal to about 20–25 Å. In Fig. 17 we report the water molecules present in the major groove, considering not only "groove" water, but also first

TABLE XI
"Grooves" Water Molecules in the Minor and Major Grooves

Limits	N_w	E^a	Groove	Solute
$4 < R < 8$	0.38	−125.38	Minor	Na^+-B-DNA
$4 < R < 8$	15.82	−104.25	Major	Na^+-B-DNA
$4 < R < 8$	0.0	0.0	Minor	B-DNA
$4 < R < 8$	9.22	−66.93	Major	B-DNA
$8 < R < 12$	19.70	−98.52	Minor	Na^+-B-DNA
$8 < R < 12$	74.02	−97.22	Major	Na^+-B-DNA
$8 < R < 12$	18.94	−72.47	Minor	B-DNA
$8 < R < 12$	64.87	−66.77	Major	B-DNA
$12 < R < R$ (max)	49.77	−88.54	Minor	Na^+-B-DNA
$12 < R < R$ (max)	64.41	−82.58	Major	Na^+-B-DNA
$12 < R < R$ (max)	64.92	−63.75	Minor	B-DNA
$12 < R < R$ (max)	102.95	−61.76	Major	B-DNA

[a] Energy in kJ/mol.

TABLE XII
Number of Water Molecules in the Grooves, N_w, and Interaction Energies, E

Limits	N_w	E^a	Groove	Solute
$0 < R < 4$	0.0	0.0	Minor	Na^+-B-DNA
$0 < R < 4$	1.15	−125.25	Major	Na^+-B-DNA
$0 < R < 4$	0.0	0.0	Minor	B-DNA
$0 < R < 4$	1.92	−87.10	Major	B-DNA
$4 < R < 8$	20.25	−137.12	Minor	Na^+-B-DNA
$4 < R < 8$	47.56	−118.48	Major	Na^+-B-DNA
$4 < R < 8$	19.79	−94.56	Minor	B-DNA
$4 < R < 8$	39.99	−81.40	Major	B-DNA
$8 < R < 12$	80.83	−117.46	Minor	Na^+-B-DNA
$8 < R < 12$	142.15	−116.88	Major	Na^+-B-DNA
$8 < R < 12$	74.36	−87.54	Minor	B-DNA
$8 < R < 12$	121.26	−87.74	Major	B-DNA
$12 < R < R$ (max)	65.56	−98.90	Minor	Na^+-B-DNA
$12 < R < R$ (max)	89.52	−96.02	Major	Na^+-B-DNA
$12 < R < R$ (max)	72.56	−68.33	Minor	B-DNA
$12 < R < R$ (max)	117.12	−67.96	Major	B-DNA

[a] Energy in kJ/mol.

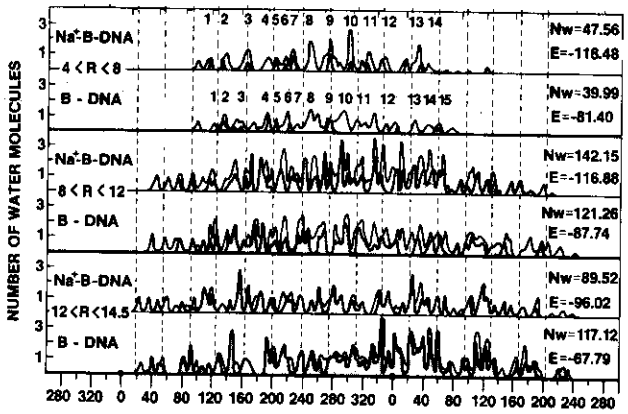


Fig. 17. Probability distributions for water molecules in B-DNA and Na^+-B-DNA major grooves: both first solvation shell and groove water molecules are considered.

solvation shell water molecules. The peaks presented as dashed areas correspond to *groove* water molecules; those presented as a line contour correspond to *groove* and *first solvation* shell water molecules. To facilitate the understanding of the figure, some of the peaks are labeled with numerals. In the first subvolumes ($0 < R < 4.0$ Å), most of the water molecules are present as first solvation shell rather than as groove waters. Peaks 1–3, 5–7, and 9–12 have a groove's water contribution, at times very small (for example, peak number 10); no *groove* contribution is sometime present (for example, peaks 4, 8, and 13). In Na^+-B-DNA the groove contribution is larger than in B-DNA because of the ion compression effect. The NH_2 groups for A and C are in the vicinity of peaks 1–3, 5, 8, 9, and 13–15; the O4(T) and the O6(G) atoms are in the vicinity of peaks 4, 5, 7, 11, and 12; and nitrogen lone-pairs (N7 of G and A) are in the vicinity of the peaks 4, 6, 8, 9, 10, 12, and 13. Rotations of water molecules (and thermal effects) are expected to be appreciable mainly for the water molecules corresponding to the dashed area.

In the subvolumes from $R = 4$ to 8 Å, each first solvation peak is enhanced by groove water contributions (exceptions are at the boundaries). The peak intensity is higher for Na^+-B-DNA than for B-DNA, as expected. The emerging picture is that a water filament starts as bound water, continues as groove water, and ends as bound water, in agreement with the previous findings.

In the subvolumes from $R = 12$ to 14.5 Å, the groove waters are responsible for most of the intensity of the peaks, especially in Na^+-B-DNA.

WATER SOLVATION OF DNA AT LOW AND MEDIUM RELATIVE HUMIDITY

In a previous paper[1] we analyzed aspects of the water organization at low and intermediate relative humidity. The simulation, however, cannot be tested with experimental data, since we selected a sample of B-DNA without counterions. (For this reason, however, the simulations on B-DNA have been performed after testing our method and our atom–atom pair potentials with examples where experimental data are available.[4,13]) As is known (see Refs. 8 and 22), in adsorption–desorption experiments, the isotherm presents a characteristic growth curve. The experiments are generally performed on DNA fibers, and in these conditions there are several parameters to keep in mind, even at constant temperature, relative humidity, and ionic concentration. By decreasing (or increasing) the number of water molecules, there is a reorganization in the entire water system. In addition, the relative position of the DNA units in the fiber varies (a swallowing of the fiber is observed). Moreover, concomitant with the reorganization of the water and of the DNA units within the fiber, conformational variation in the DNA can occur. *Finally*, fiber diffraction data are limited to about 3-Å resolution. An advantage in simulating the water structure around a fixed-geometry Na^+-B-DNA fragment is that the energetic and structural variations of the water system can be followed, without the worry of the notable additional complications associated with the interactions of one DNA molecule (and its solvent) with other DNA molecules in the fiber.

We precede the following analyses by reporting the average energies computed by adding either 22, 44, 142, or 257 water molecules to the Na^+-B-DNA fragment. The results are reported in Tables XIII–XVI and in Figs. 18 and 19. A simulation performed with 371 water molecules was performed but is not analyzed because it yields results very similar to those analyzed for the case of 447 water molecules.

For 22 waters the sites of interest are Na^+, O1P, and O2P; one water molecule is bound to each PO_4^-–Na^+ unit. One might have expected that all the 22 waters are bound only to the Na^+ ions (or to either O1P or O2P); but from Table XIII one learns instead that a more involved pattern exists. A water molecule is bound with about equal probability to Na^+ and O2P but less preferentially to O1P. The O5′ and O3′ sites "do not participate" (the analyses on O1P and O2P yield a sum of 17.80 water molecules (Tables XIII and XIV); also considering O5′ and O3′ and the CH_2 group does not change this value. Essentially, two distributions are found, one with a water bound both to Na^+ and O2P, and a second, less probable one with a water bound both to Na^+ and O1P. The large water population counted in the first solvation shell for the hydrogens H3 and H4 of the sugar atoms is simply a *recounting* of the above water molecules (bound to PO_4^-–Na^+), which can also be considered *as if* in the first hydration sphere of the sugar atoms (shared area in Fig. 4).

TABLE XIII
Medium- and Low-Humidity Distributions for Na^+-B-DNA at 300 K: First Solvation Shell (Site Analyses)[a]

No.	Site	142 Water Molecules				44 Water Molecules				22 Water Molecules			
		Bound		Complement		Bound		Complement		Bound		Complement	
1	Na^+	56.39	−134.98	0.0		22.70	−148.81	0.0	—	13.37	−156.89	0.0	—
2	O1P	41.50	−133.47	11.56	−126.31	14.84	−146.02	0.37	−133.17	5.64	−152.14	0.02	−148.22
3	O2P	35.48	−140.34	9.89	−137.67	18.05	−152.38	1.37	−144.28	12.16	−158.48	0.02	−142.21
4	O5′	0.69	−123.68	0.0	—	0.0	—	0.0	—	0.0	—	0.0	—
5	O3′	12.82	−115.60	1.05	−134.27	0.18	−143.95	0.07	−140.00	0.02	−146.10	0.0	—
6	O1′	1.19	−130.83	0.0	—	0.0	—	0.0	—	0.0	—	0.0	—
7	NH_2-G	0.13	−146.34	0.25	−140.67	0.0	—	0.0	—	0.0	—	0.0	—
8	O2-T	0.54	−131.73	0.0	—	0.0	—	0.0	—	0.0	—	0.0	—
9	O2-G	0.22	−148.11	0.0	—	0.0	—	0.0	—	0.0	—	0.0	—
10	CH_2 at PO_4^-	—	—	18.01	−126.28	—	—	1.36	−142.58	—	—	0.20	−142.97
11	H3, H4 of S	—	—	41.63	−139.79	—	—	19.43	−148.96	—	—	11.50	−155.60
12	H2's of S	—	—	24.26	−140.02	—	—	5.64	−147.27	—	—	1.63	−153.38
13	H1 of S	—	—	0.85	−135.16	—	—	0.0	—	—	—	0.0	—
14	H's of T	—	—	5.90	−132.38	—	—	0.0	—	—	—	0.0	—
15	H's of A	—	—	1.52	−125.07	—	—	0.09	−141.98	—	—	0.0	—
16	H's of C	—	—	3.33	−131.41	—	—	0.0	—	—	—	0.0	—
17	H's of G	—	—	4.54	−138.85	—	—	0.38	−139.44	—	—	0.0	—
18	I-Boundary	11.26	−128.78	2.49	−128.81	4.73	−137.83	0.42	−125.25	2.25	−150.13	0.02	−138.05
19	II-Boundary	7.56	−128.28	0.04	−106.19	4.20	−137.72	0.0	—	3.04	−146.65	0.0	—
20	H-Boundary	—	—	5.97	−124.72	—	—	0.77	−129.41	—	—	0.02	−148.69
21	*Total*	140.57	−133.30	—	—	44.00	−146.88	—	—	22.00	−154.41	—	—

[a] The sum of the "bound" and of the "complement" water constitutes the first hydration shell. The boundaries are formed by atoms of either type 1 or type 2 (hydrogen-bonded to a water molecule) or by hydrogen atoms.

TABLE XIV
Medium- and Low-Humidity Distributions for Na^+-B-DNA at 300 K: First Hydration Shell, Groups of Sites

No.	Groups of Sites[a]	142 Water Molecules				44 Water Molecules				22 Water Molecules			
		Bound		Complement		Bound		Complement		Bound		Complement	
1	2,3,4,5,10	85.16	−134.51	23.36	−131.32	32.69	−149.53	1.63	−142.64	17.72	−156.47	0.02	−139.60
2	2,3,4,5,10,1	121.53	−134.48	2.40	−118.55	38.02	−148.45	0.0	—	18.82	−155.75	0.0	—
3	Sugars	1.19	−130.83	59.22	−139.17	0.0	—	21.91	−149.11	0.0	—	11.85	−156.06
4	Bases-sugars	1.65	−131.28	61.26	−138.93	0.0	—	22.20	−148.98	0.0	—	11.85	−156.06
5	A	0.0	—	1.52	−125.55	0.0	—	0.09	−142.48	0.0	—	0.0	—
6	C	0.22	−148.11	3.33	−131.88	0.0	—	0.0	—	0.0	—	0.0	—
7	G	0.13	−146.34	4.79	−139.42	0.0	—	0.38	−139.92	0.0	—	0.0	—
8	T	0.54	−131.73	5.90	−132.86	0.0	—	0.0	—	0.0	—	0.0	—
9	A·T	0.54	−131.73	7.04	−131.27	0.0	—	0.9	−142.48	0.0	—	0.0	—
10	G·C	0.23	−147.38	8.01	−136.18	0.0	—	0.38	−139.92	0.0	—	0.0	—
11	A,T,G,C	0.77	−136.46	14.10	−133.62	0.0	—	0.47	−140.39	0.0	—	0.0	—
12	Grooves	1.43		−96.47		0.0		—		0.0		—	
13	First shell	140.57		−133.40		44.0		−146.88		22.0		−154.41	
14	Total (W-W)	−4.64 ± 0.13				−1.81 ± 0.10				−1.92 ± 0.11			
15	Total (W-S)	−132.23 ± 0.29				−145.12 ± 0.27				−152.50 ± 0.25			
16	No. conformation[b]	120000				80000				20000			

[a] The numerals in this column identify the sites included in a group; each numerical corresponds to the number of the first column of Table XIII.
[b] In the analysis for the data reported here and in Table XIII (not in the Monte Carlo simulations). See Tables I–IV for equivalent tables at high humidity.

TABLE XV
Solvation of Na^+-B-DNA at 300 K with 257 Water Molecules: "Bound" Water[a]

No.	Site Name	Waters	Energy	Group of Sites	Waters	Energy
1	O1P	34.12	−140.80	1,2,3,4	77.44	−140.34
2	O2P	35.94	−144.63	1,2,3,4,5	107.94	−140.06
3	O5′	1.00	−134.45	7,8,9	22.75	−128.64
4	O3′	10.39	−121.21	10,11,12,13	8.47	−131.94
5	Na^+	46.22	−141.18	14,15,16,17	15.34	−132.00
6	O1′	15.87	−138.57	A	9.25	−128.81
7	NH_2 (A)	6.30	−125.97	C	13.96	−124.92
8	NH_2 (C)	11.22	−122.46	G	16.73	−132.17
9	HN_2 (G)	7.86	−139.18	T	9.69	−132.04
10	N3 (A)	1.60	−134.06	A-T	16.37	−131.23
11	N7 (A)	2.09	−135.47	G-C	26.43	−129.49
12	N3 (G)	2.23	−135.00	A-T, G-C	32.37	−129.92
13	N7 (G)	2.54	−125.02	A-T, G-C,6	36.40	−130.38
14	O2 (T)	3.35	−145.06			
15	O4 (T)	6.34	−125.17			
16	O2 (C)	2.74	−135.01			
17	O6 (G)	5.63	−124.70			
18	Boundary (type 1)	9.04	−133.23			
19	Boundary (type 2)	5.16	−137.26			

[a] Energy in kJ/mol.

TABLE XVI
Solvation of Na^+-B-DNA at 300 K with 257 Water Molecules: "First Solvation" Shell[a]

No.	Site Name	Waters	Energy	Group of Sites	Waters	Energy
1	O1P	40.73	−139.40	1,2,3,4,18	117.74	−136.90
2	O2P	46.45	−143.98	1,2,3,4,5,18	133.86	−137.11
3	O5′	1.00	−134.45	7,8,9	26.31	−128.17
4	O3′	11.01	−122.22	10,11,12,13	8.89	−130.63
5	Na^+	46.45	−141.14	14,15,16,17	15.35	−132.00
6	O1′	15.91	−138.51	A	16.24	−126.96
7	NH_2 (A)	8.31	−124.91	C	21.57	−124.51
8	NH_2 (C)	13.13	−122.28	G	22.67	−132.70
9	NH_2 (G)	9.12	−138.23	T	21.30	−130.30
10	N3 (A)	1.77	−132.51	A-T	31.45	−129.99
11	N7 (A)	2.18	−133.78	G-C	38.56	−129.22
12	N3 (G)	2.23	−135.00	A-T, G-C	56.21	−129.30
13	N7 (G)	2.71	−122.78	A-T, G-C, S	126.66	−133.16
14	O2 (T)	3.45	−144.52	Sugar	85.25	−135.56
15	O4 (T)	6.76	−124.39	Grooves	71.32	−94.81
16	O2 (C)	2.74	−135.01	First shell	185.69	−132.55
17	O6 (G)	5.63	−124.70			
18	CH_2 at PO_4^-	44.21	−127.14	Total interaction[b]	−122.8 ± 0.07	
19	H3, H4, of S	65.00	−135.20	Water–water[b]	−17.77 ± 0.02	
20	H2's of S	28.17	−138.92	Water–DNA	−105.11 ± 0.05	
21	H1 of S	10.14	−137.08			
22	H's of T	14.59	−128.29			
23	H's of A	5.17	−126.28			
24	H's of C	9.21	−122.63			
25	H's of G	4.80	−135.55			
26	I-Boundary	10.45	−134.32			
27	II-Boundary	5.16	−137.26			
28	H-Boundary	8.64	−120.45			

[a] Energy in kJ/mol.
[b] For the full set of 257 water molecules.

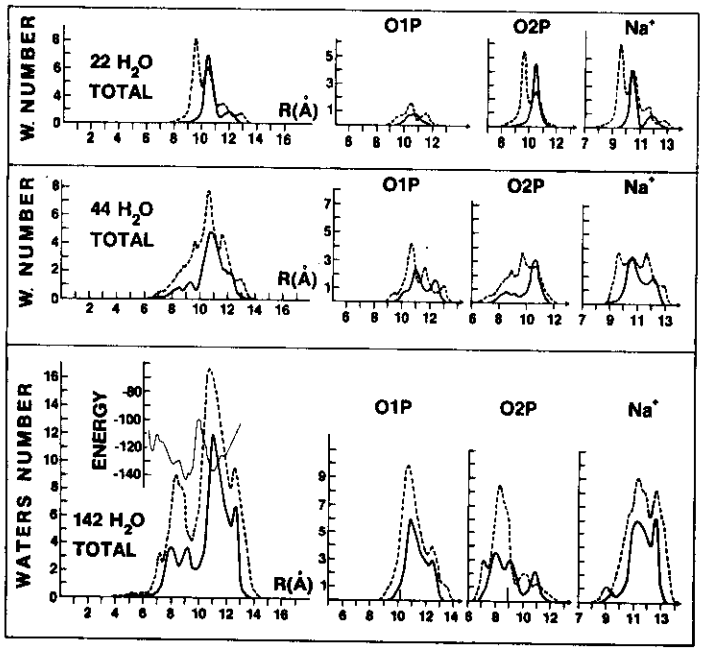


Fig. 18. Probability distributions for 22, 44, and 142 water molecules in Na^+-B-DNA at 300 K (see Figs. 2 and 3).

In order to accommodate 22 *additional* water molecules, the original distribution for the 22 water molecules is fully rearranged. Indeed, for the case of 44 water molecules, the O3′ site starts to become involved; the O1P and the O2P site's difference, in water population, becomes smaller than for the case of 22 water molecules; and the average energy of a water molecule interacting with PO_4^-–Na^+ drops from −155.75 kJ/mol (for 22 waters) to −148.45 kJ/mol. The rearrangement pushes the water molecules (in average) to smaller R values, and not only the hydrogen atoms of sugar but also those of the A and G bases are becoming involved, even if marginally.

Let us now fill the DNA sample with 142 water molecules; the grooves remain essentially empty (1.43 water molecules out of 142 are found in the grooves), but the finding of some, even if with very little, indicates that the value for "bound" water has attained saturation. This statement is not equivalent to saying that the structure or the maximum numbr of the *bound* water molecules for 142 water and for 447 water molecules solvating the DNA fragment is the same. Indeed, by filling up the grooves, additional water penetrates the "bound" region, and because of the interaction between water in the "grooves" and "bound" water, a rearrangement can be

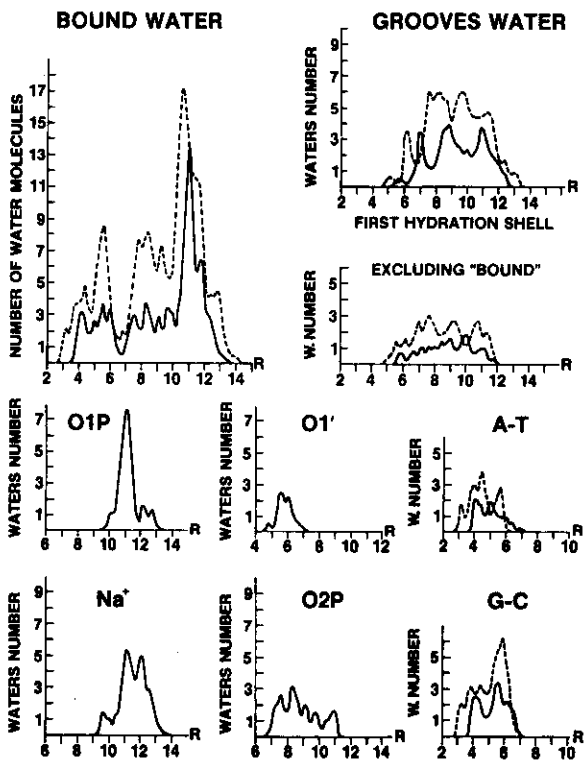


Fig. 19. Probability distributions for 257 water molecules in Na^+-B-DNA at 300 K (see Fig. 17).

induced in the "bound water" structure. In other words, we are gathering preliminary evidences that the adsorption process is not "monotonic"; at low relative humidity the water molecules go mainly to the DNA sites; at higher humidity, mainly to the grooves; and at even higher values to the sites and the grooves and then, finally, to the "extended" grooves. Incidentally, it should be evident by now why the interpretation of experiments at different humidities (desorption–adsorption studies) is very hard (even neglecting the experimental difficulty in obtaining reproducible raw data), especially since they are often performed on DNA fibers, where additional variables must be considered. For the case of 142 waters, about 2.97, 2.18, and 1.87 water molecules are bound to Na^+, O1P, and O2P, respectively; the total populations of *bound* water for PO_4^- and PO_4^-–Na^+ are 4.48 and 6.40 water molecules, respectively. The total populations for the *first hydration shells* for PO_4^- and PO_4^-–Na^+ are 5.71 and 6.52, respectively. The previously noted drop in the binding energy for the "bound" water

molecules around PO_4^-–Na^+ continues, and the energy is now −134.48 kJ/mol. The bases are involved in the solvation, and the solvation order is A < G < C < T when considering the number of "bound water" molecules, A < C < G < T when considering the number of waters in the first solvation shell, and A < C < T < G when considering the binding energy for the water molecules in the first solvation shell. The distribution of the total water population at different R values shows a well-defined structure (see Fig. 18); by comparing the shift of the most intense peak, one gains a feeling for the intense water structure rearrangement that follows the repeated addition of water molecules. The A-T pair is more solvated than the G-C pair for the cases with 44 and 142 water molecules, in agreement with experimental findings at low humidity.[22] In addition, we note that at low humidity, O2P is more solvated than O1P, a situation that is reversed by increasing the relative humidity. Finally, in presence of Na^+ counterions, the O1P and the O2P water distributions are nearly additive, contrary to what is found for B-DNA without counterions.

Let us now solvate the DNA fragment with 257 water molecules (see Fig. 19 and Tables XV and XVI). This case deserves to be analyzed more deeply, since it provides an interesting step in the adsorption process. With 257 water molecules, the first hydration shell is nearly saturated, and additionally, there are about 71 water molecules in the grooves. From Fig. 19 we learn that the bound water molecules exhibit the same 10 peaks as found in the 447 water molecules case; however, the peaks L and J (see Fig. 3) are more developed in adding water to the 257 molecules now present, because the Na^+ and the O1P sites are not fully solvated. With 257 water molecules, the hydration at the base pairs is evolved nearly as much as with 447 water molecules (peak C is, however, higher than in the case of 447 water molecules). The water molecules in the "classical" grooves are well developed at smaller R values but far from saturation at larger R values, as expected. Therefore, the term "*groove*" water could be substituted by the term "*second solvation shell*" for the case of 257 waters. Relative hydrations at the bases are not discussed, since the data of Table XVI are self-explanatory.

In a study in progress, we are analyzing the differences in rotational freedom associated to the water molecules either bound or in the grooves at different relative humidity. This is an aspect of interest for the interpretation of angular distribution data from neutron scattering. We note that the findings concerning the water filaments are in agreement with the scattering data by Dahlborg and Rupprecht.[23] In considering the partial humidity simulations above, the following model for an adsorption process emerges. First, the water molecules go to the phosphate groups. If the humidity is very low [one water molecule per phosphate (22 water case)], then a given water has several nearly equivalent choices of position at each phosphate and the *same* position will not be selected. By an increase in the humidity, such that two water molecules are available for each phosphate group (44 water case), the best solvation positions are *not* those left

unoccupied by the first water molecule, but new ones. This is to be expected: the Na^+–PO_4^- system is different from the H_2O–Na^+–PO_4^- system, and the latter has only a vague "memory" of the former. Increasing humidity such that, on average, about 6.5 water molecules are available for each phosphate group (142 waters case), all waters solvate the Na^+–PO_4^- groups, but some molecules solvate even the base pairs or are present in the groove region. The presence of water in the groove indicates that a saturation level has been reached. Alternatively, we can state that the initial saturation of the grooves is nearly concomitant to the solvation at the base pairs. By an increase in humidity such that about 11.7 water molecules are available for each phosphate group (case of 257 waters), more water molecules go to the groove region and the solvation of the sugar oxygen and of the base pairs is nearly completed. Finally, by making available about 20.2 water molecules per phosphate group (447 waters), additional water is packed into the groove region and around the solvated sites, but most of it goes at the outside of Na^+-B-DNA, the "extended groove" region, still strongly perturbed by the Na^+-B-DNA field, so much so that the bulk water interaction is far from being reached. By adding two additional water layers (that is, by adding 1600 water molecules), one would have nearly reached the bulk water region, at an R value of about 20 Å. This model holds at a temperature of 300 K; at lower temperature, a higher number of water molecules will solvate Na^+-B-DNA, due to the decrease of thermal disorder.

CONFORMATIONAL TRANSITIONS: A MODEL AND A PRELIMINARY EXAMPLE

In considering the adsorption process, we should keep in mind some overall energetic aspects, reported in Fig. 20. Repeatedly, we have pointed out that the interaction of water with DNA increases by adding counterions to the solute, which therefore stabilizes Na^+-B-DNA relative to B-DNA. This increase is drastic, as shown in Fig. 20, where the total (namely, the sum of water–DNA and water–water) average interaction energy per water molecule is reported as a function of the number of water molecules, N_w, solvating the DNA fragment. In the top insert of Fig. 20 the full line refers to the *total* interaction, the dashed line to the *water–DNA* interaction; the *water–water* interaction is reported in the bottom insert. By a substantial increase of solvent all curves will eventually end at about −36 kJ/mol, the bulk water value. The total interaction might suggest that the adsorption process is monotonic; the partial interactions offer indications for an opposite conclusion. Physically, the following model clearly emerges *in B-DNA*: the loss of bulk water energy (water–water interaction) is more than compensated by the strong PO_4^- attraction; in addition, the water–water interaction at low humidity would amount to only a fraction of the bulk-water value. But when the PO_4^- groups are screened by few solute molecules, and when the humidity increases, then the system attempts to regain

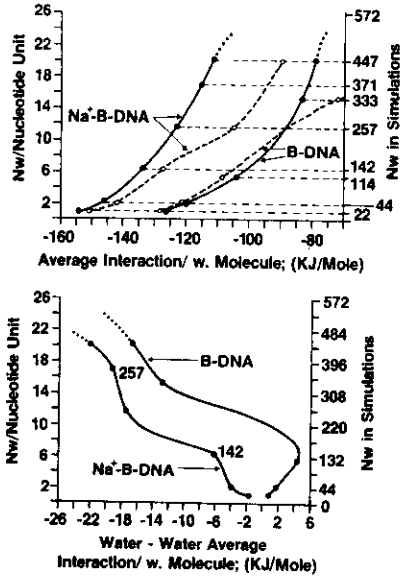


Fig. 20. Average interaction energies per water molecules (in kJ/mol) during adsorption. Top insert: Total interactions (solid line), water–B-DNA interactions (dashed lines). Bottom insert: water–water interaction.

the bulk-water interaction energy since the water–water interaction becomes more and more important. *In Na^+-B-DNA* (and in any B-DNA counterions system), the initial adsorption situation is more complex. At low humidity, a water molecule does not experience only the PO_4^- field, but also the counterion field. The latter exerts on a water molecule an orienting field yielding a very different orientation from the one originated by the PO_4^- field. Thus, the orientation of the water molecule is the global response to the two fields; the resulting orientation is not as repulsive to another water molecule as the one obtained for the case of B-DNA (without counterions and at low humidity).

As a consequence of these effects, the water–water interaction energy presents two plateaus, where a small humidity increase brings about a strong water–water interaction variation. The two plateaus border at two steep sides, where a large humidity increase brings about only a small water–water interaction variation.

The exact shape of the curves we have drawn is limited by the number of our computer experiments; we would not be too surprised if the plateau between the two simulations with 142 molecules and 257 molecules would turn out to be flatter than in our graph. The relative humidity interval (corresponding to a variation from about 6 to about 10 water molecules per nucleotide unit) would appear as a *critical humidity interval.* The classical

grooves start to be filled in this interval; by a further increase in the humidity, the final saturation at the DNA solvated sites is reached and the "extended" grooves are filled. When the adsorption process is completed (very high humidity), it is hard to visualize a conformational transition in DNA (for example, from the conformation A to B) without the introduction of a third body. Equally difficult is visualizing a conformational transition at very low humidity, since the few water molecules present are fully occupied at solvating the PO_4^- counterion groups. On the other hand, in the above-discussed plateau, where the incipient filling up of the classical grooves perturbs the sugar units (this is tantamount to a directional pressure), one can easily hypothesize conformational transitions. Thus, combining the evidence regarding the ion-induced compression effect and the results on the variations of the interaction energies with relative humidity, we can obtain a preliminary model to explain conformational transition, under specific conditions. The use of the model, however, requires simulations of the type here presented for different conformations and with different counterions. The mathematical conditions for a conformational transition are now indicated by the following three models.

First, we consider only one type of counterion (for example, Na^+). Let $E(A)$ and $E(B)$ be the energy of DNA in conformations A and B, respectively, for DNA fragments of equal number of units. From a simulation with an equal number of water molecules and at a given temperature, one can obtain the water–DNA interaction $ES[A,h(i)]$ and $ES[B,h(j)]$, where the indices A and B refer to the two conformations, and $h(i)$ and $h(j)$ to the two relative humidities (clearly different in the two conformations, since we have assumed the same number of water molecules). A transition can occur for $\Delta E = \Delta E'$, where $\Delta E' = E(A) - E(B)$ and $\Delta E = ES[A,h(i)] - ES[B,h(j)]$, *assuming about equal entropic variations.* With this work we provide one of the four needed quantities.

Second, we consider the same situation as above, with, however, two different types of ions $I(1)$ and $I(2)$ (for example, Na^+ and Li^+) yielding either $I(1)$–DNA or $I(2)$–DNA. Let $E[A,I(1)]$, $E[A,I(2)]$, $E[B,I(1)]$, and $E[B,I(2)]$ be the energies of the fragment with either one of the two types of counterions and in the two conformations. Correspondingly, let $ES[A,I(1)]$, ..., $ES[B,I(2)]$ be the corresponding simulated solvation energies as in the previous model. The introduction of more than one type of counterion increases the probability of finding an equality between ΔE and $\Delta E'$ since one more variable, the type of ion, has been added.

Third, let us assume the same situation as above, with, however, the additional possibility that a fractional rather than total replacement of $I(i)$ with $I(j)$ is considered. This last assumption considerably increases the probability of finding an equality between ΔE and $\Delta E'$. The model can be made more complete by including a temperature index, namely, by considering more than one temperature. The importance of temperature variations in discussing the stabilization of DNA due to the solvent has been previously stressed; see, for example, Lewin.[12] As an example, let us con-

sider the *ES* values, stressing, however, that simulations on Li^+–B-DNA, Li^+–A-DNA, Na^+–A-DNA are not available. Different tentative solvation energy values can be chosen by keeping in mind ionic solvation studies,[9-11] the differences in the solvation energies between the A and B single-helix conformation and preliminary simulation data on Li^+–B-DNA. We have made no use of the study by Marynick and Shaeffer,[10] since their use of a subminimal basis set and the neglect of basis set super-position corrections deprive their otherwise very interesting computation of quantitative validity. Their very strong attraction for a phosphate group in $(CH_3\text{-}O)_2\text{-}PO_2^-$ with Li^+ and Na^+ (computed as −303 and −198 kcal/mol, respectively) and the very short oxygen–cation distances (computed as 1.77 and 1.98 Å, respectively) are too far from the more reasonable data obtained, however, on $(CH_3\text{-}CH_2\text{-}O)_2PO_2^-$ and yielding −152 and −134 kcal/mol, and 1.93 and 2.17 Å, respectively (manuscript in preparation). Equivalent data, obtained from relatively old electrophoretic mobility studies, indicate that the binding to DNA for ions is $Li^+ > Na^+ > K^+$.[24]

At low humidity (1 water per nucleotide unit), Na^+–A-DNA is estimated to be more attractive to a water molecule by about 8 kJ/mol (relative to A-DNA). Simulations on A and B single helix[2,3] indicate a value between 5 kJ/mol (average value) and 17 kJ/mol (maximum difference); thus our selected value might be somewhat smaller than the correct one. At high humidity we know that the PO_4^- groups in A-DNA are sufficiently crowded relative to B-DNA such that they have one less water molecule strongly bound to a PO_4^- group. Thus, on a sample of 447 water molecules, about 22 water molecules in A-DNA are less bound to the PO_4^- than in B-DNA by an assumed interaction of about 10 kJ/mol per water molecule. As a result, the total interaction energy curve for Na^+-A-DNA crosses the total interaction energy curve of Na^+–B-DNA at a relative humidity corresponding to about 160–190 water molecules (see Fig. 20). For Li^+-B-DNA, we have obtained preliminary data for 22 and 447 water molecules (at a temperature of 300 K). At a relative humidity corresponding to 22 water molecules, Li^+-B-DNA attracts more water than Na^+-B-DNA, but the situation is reversed at high humidity (447 water molecules); this brings about a crossing of the Li^+-B-DNA and Na^+-B-DNA total interaction energy curves at about 190–210 water molecules. Finally, for Li^+-A-DNA, we assume that the total interaction energy to water differs from the Na^+-A-DNA total interaction energy in the same way as found by comparing Li^+-B-DNA with Na^+-B-DNA. Until definitive simulations on Li^+-A-DNA, Li^+-B-DNA and Na^+-A-DNA are available, the above estimates are probably all the data one can use. The stabilizations due to the solvent effect in the A → B conformational transition are reported on Fig. 21 (the ordinate gives the number of water molecules for either an A or a B double helix with 22 phosphate units). We consider four cases (all at 300 K). In case 1, we consider the stabilization of a DNA conformation with Na^+ counterions, whereas in case 2, we consider the stabilization of a DNA conformation with Li^+ counterions. In case 1, at low humidity, form A is

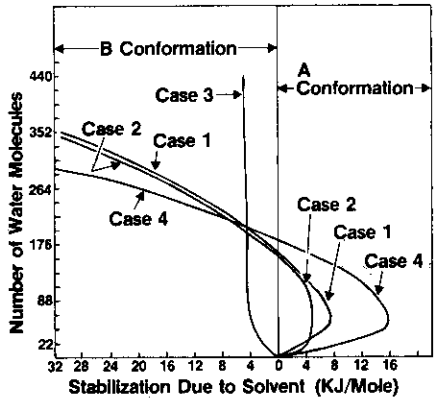


Fig. 21. Solvent stabilizations at constant temperature (300 K) and different humidities for the A → B transition; case 1 refers to Na^+-A-DNA → Na^+-B-DNA, case 2 to Li^+-A-DNA → Li^+-B-DNA, case 3 to Na^+-A-DNA → Li^+-B-DNA, and case 4 to Li^+-A-DNA → Na^+-B-DNA (see text).

stabilized by water, reaching a maximum for about 3 water molecules per nucleotide unit and then going to zero at about 7 water per nucleotide unit; then the B form becomes stabilized. In case 2 (Li^+ counterions), the same behavior is predicted, but the crossing from A to B occurs at slightly lower humidity. In cases 3 and 4, we increase not only the humidity, but we also assume that the Na^+ counterions of the A form are substituted with Li^+ counterions in the B form (case 3) or vice versa (case 4). From our preliminary data we expect that only form B is stabilized by the solvent, whereas in case 4, the Na^+-A-DNA has a net solvation stabilization up to about 7 water molecules per nucleotide unit; at higher humidity the solvent effect helps the formation of B-DNA.

We stress that we have referred only to ΔES, not to $\Delta E'$; in addition, no entropic effect has been considered, or equivalently, we have assumed that the entropic contribution to the free energy is ion-independent and conformation-independent at a given relative humidity and temperature. The theoretical behavior of Fig. 21, even keeping in mind its tentative nature, explains a large number of experimental findings relative to the A–B transition. Clearly, the same type of reasoning can be used when considering the solvent effect of any other conformational transition. By adding to an ionic solution (containing A- or B-DNA) solvents like alcohol–water, the number of water molecules available to DNA decreases because the hydrophobic part of the alcohol removes water from DNA.[25,26] Thus, if one can estimate the latter effect, then Fig. 21 provides an explanation, also for transitions, where not only the humidity and the counterions are varied, but also additional solvents like alcohols are added. Concerning the energy $\Delta E'$ for conformational transitions from A to B, a value of about 84 kJ/mol

has been proposed by Ivanov and coworkers[27,28] and later confirmed by and Sukhorukov and coworkers.[29]

CONCLUSION

These simulations of the water structure around Na^+-B-DNA and B-DNA are likely the most thorough computational attempt up to now available. The tables we have provided and the detailed discussions and figures allow reinterpretation of experimental data. Unfortunately, we are not in a position to carry out this task, since it would require access to the *raw* experimental data. In the event our data will be used to interpret fiber data, we stress that the simulations presented here deal with a single DNA unit, and our predictions should therefore be used with care, especially for large R values. It is likely that our findings for R values up to about 9 Å can be used directly; those from R = 9 to 12 Å only to provide an analogy; and in the case of those from R = 12 to 14.5 Å, the water structure in a fiber is expected to be quite different from the one here presented.

A practical approach could be as follows. Let R' be the DNA unit-to-unit distance in a fiber (at a given temperature and relative humidity). The corresponding probabilities and/or energy graphs of our study can be used, placing our diagram origin for R at the long axis of the two DNA units (the same graph is used twice as two mirror images, with the symmetry plane at $R'/2$). The resulting overlap region near $R'/2$ should be disregarded, and the same holds for about 3 Å on both sides of $R'/2$. (Namely, one should be careful to disregard an interface region of about 2 water molecules, one at each side of $R'/2$.) We note that a simulation on a minifiber is technically feasible; the main problem would be the long computer time involved, the determination of the rotation of one unit relative to the other and the determination of the counterion position, which might be different at the surface, relative to the inner part of the minifiber.

Strictly speaking, the data predicted here should be compared with B-DNA in solution at 300 K. The notable scarcity of such data will likely be complemented by future neutron-beam experiments, actively pursued at a number of laboratories, but still at a very early stage.

There are several approximations in our treatment, which limit the realism of our simulation. The Na^+–B-DNA fragment we have used is rather small, only 782 atoms. We do not expect, however, that there are gross errors in our analyses, since we have considered the top and the bottom of our fragment separately from the main body. A larger fragment would certainly increase the computational times considerably, but it would probably yield essentially the same results. Explicit inclusion of rototranslational periodic boundary conditions could be considered, but not with the base-pair selection we have used. We have included only 447 water molecules; this is probably sufficient if one is interested in the first solvation shell and in the groove water molecules. It is noted that the addition of two more shells of water would require the inclusion of about 1600 additional water molecules. Again, the computation would be much more costly, and the net advantage (if any) is not clear.

The most crucial limitation, however, is in the assumption of a rigid Na^+-B-DNA fragment, and in the position assumed for the Na^+ ions. From theoretical considerations the lowest energy minimum (minimum a) for Na^+ in model compounds containing the phosphate group is in agreement with our choice. However, there is a second minimum away from the PO_2^- plane (minimum b) that will become deeper if more than a single phosphate group is present; finally, there is a third minimum along the PO bond direction (minimum c). The three minima are relatively close in energy. Let us consider the available, but indirect, experimental, data. A double-helix chain can be constructed from the x-ray structure of sodium adenylyl-3′,5′-uridine (ApU), where one Na^+ is coordinated to the PO_4^- group and one to the uracil carbonyl groups.[30] Thus, minimum b is known to be a possible candidate in polynucleic acids. A double-helix chain can be constructed from the x-ray structure of sodium guanylyl-3′,5′-cytidine (GpC), where the Na^+ is coordinated to the two free oxygen atoms of a phosphate group.[31] Thus, minimum a is a possible candidate. Finally, x-ray studies on crystals[32] of the deoxyribose dinucleotide sodium thymidylyl-5′,3′-thymidylate-5′ (pTpT) can be used to model a double helix with the Na^+ coordinated to one free oxygen of one PO_4^- group and two oxygen atoms on two different thymine bases. This position for Na^+ can be considered as one related to the minimum c. It is noted that none of the above structures refers directly to a crystal of a true polynucleic acid nor to a 50% G-C, 50% A-T double-helix structure, as in our fragment. The nearest case is the one of GpC, where the Na^+ is located at about our chosen position.

We note that analyses of the Monte Carlo data for a "free" Na^+ simulation would be notably more complex, since one would expect a probability distribution for the Na^+ position over a volume of few angstroms.[3] As a result, the equilibration process in the Monte Carlo simulation and the data collection would require a notable increase of computer time. These estimates are guided by a simulation of Na^+ placed in a water solution with a glycine zwitterion as solute (manuscript in preparation). We have found that the Na^+ did not remain at the energy minimum (determined by neglecting the solvent), but shifted away by about 1 Å. This shift allows the development of a more complete solvation shell both for the COO^- group and for Na^+ ion, which, however, *remains bound* to the COO^- group (namely, no interstitial water penetrates between the COO^- group and Na^+). In conclusion, a simulation with "free" Na^+ ions might bring about an increase of one additional strongly bound water molecule at each PO_4^-–Na^+ group due to a shift of Na^+ from the present position to one corresponding to the above-described minimum b. Work is in progress to determine the position of the counterions.

It is likely that new coordinates for DNA will become available directly from crystal data rather than from modeling based on small fragments.

The above-quoted studies[30–32] and the recent work by Kennard and coworkers[33] point strongly in this direction. A new simulation could consider these new sources for the coordinates, still leaving open, however, the long-standing problem concerning the degree of confidence one should place on accurate x-ray data obtained *from a crystal* when one needs accurate coordinates for B-DNA *in solution*.

Finally, we recall that the macroscopic properties of DNA in solution depend on the chain length.[34,35] However, since low-molecular-weight DNA exhibits a rigid rodlike helix behavior,[36] our simulated result should be used with care, if applied to high-molecular-weight DNA chains. The chain length is also important in determining the number of counterions either bound or very near to DNA. In this respect, in our computer experiment we might have selected an excess of one to two sodium ions.[37,38] On the other hand, this choice is forced by our approximation of assuming one counterion bound to each phosphate group. Work is in progress to link our simulation with the thermodynamic models presented by Record and coworkers[37] and by Manning.[38]

To conclude, in this work we have considerably sharpened some of the questions both experimentalists and theoreticians would like to answer concerning the structure of solvent water (at nonzero temperature) for B-DNA with counterions and presented a model for predicting conformational transitions in nucleic acids. The agreement between our simulations and a large number of experimental data is very gratifying. Work is in progress to refine the conformational transition model and to include counterions such as Li^+, K^+, Ca^{2+}, and Mg^{2+}.

Note added in proof: A study with a three-turn B-DNA fragment, 1200 water molecules, and 60 Na^+ ions with statistically determined positions in the solvent has been completed since submission of this work.[39]

One of us (G.C.) would like to thank the National Foundation for Cancer Research for a fellowship and also the IBM Corporation for making available computer time on the IBM 370/3033 at the IBM Poughkeepsie Laboratories.

References

1. Clementi, E. & Corongiu, G. (1981) *Biopolymers* **20,** 551–571.
2. Clementi, E. & Corongiu, G. (1979) *Int. J. Quant. Chem.* **16,** 897–915.
3. Clementi, E. & Corongiu, G. (1979) *Biopolymers* **18,** 2431–2450.
4. Clementi, E. (1980) *Lecture Notes in Chemistry, Vol. 19, Computational Aspects for Large Chemical Systems*, Springer-Verlag, Berlin.
5. Clementi, E., Corongiu, G. & Lelj, F. (1979) *J. Chem. Phys.* **70,** 3726–3729, and references therein.
6. Clementi, E., Scardamaglia, R. & Cavallone, F. (1977) *J. Am. Chem. Soc.* **99,** 5531–5544.
7. Feldman, R. (1976) *Atlas of Macromolecules*, Document 13.2.1.1.1, NIH, Bethesda, Maryland.
8. Falk, M., Hartman, K. A. & Lord, R. C. (1964) *J. Am. Chem. Soc.* **85,** 391–398.
9. Barsotti, R. & Clementi, E. (1977) *Theor. Chim. Acta* **42,** 101–120.

10. Marynick, D. A. & Schaffer, H. F., III (1975) *Proc. Natl. Acad. Sci. USA*, 3794–3798.
11. Clementi, E. & Barsotti, R. (1978) *Chem. Phys. Lett.* **59,** 21–25.
12. Lewin, S. (1976) *J. Theor. Biol.* **17,** 181–212.
13. Clementi, E. (1976) *Determination of the Water Structure and Coordination Numbers for Ions, Lecture Notes in Chemistry*, Vol. 2, Springer-Verlag, Berlin.
14. Clementi, E. & Corongiu, G. (1979) *Chem. Phys. Lett.* **360,** 175–178.
15. Clementi, E. & Corongiu, G. (1979) *Gazz. Chim. Ital.* **109,** 201–205.
16. Clementi, E. & Corongiu, G. (1979) *J. Chem. Phys.* **72,** 3979–3992.
17. Pullman, A. & Perahia, D. (1978) *Theor. Chim. Acta* **48,** 29–36.
18. Pullman, A., Pullman, B. & Bethod, H. (1978) *Theor. Chim. Acta* **47,** 175–192.
19. Perahia, D. & Pullman, B. (1977) *Biochim. Biophys. Acta* **474,** 349–362.
20. Ranghino, G. & Clementi, E. (1978) *Gazz. Chim. Ital.* **109,** 170–189.
21. Neidel, S., Berman, H. M. & Shieh, H. S. (1980) *Nature* **288,** 129–133.
22. Texter J. (1978) *Prog. Biophys. Mol. Biol.* **33,** 83–98.
23. Dahlborg, U. & Rupprecht, A. (1977) *Biopolymers* **10,** 849–863.
24. Ross, P. D. & Scruggs, R. L. (1964) *Biopolymers* **2,** 79–89.
25. Frisman, E. V., Slonitsky, S. V. & Vaselkov, A. N. (1979) *Int. J. Quant. Chem.* **16,** 847–855.
26. Frisman, E. V., Vaselkov, A. N., Solnitsky, S. V., Karavaev, L. S. & Vorob'ev, V. E. (1974) *Biopolymers* **13,** 2169–2178.
27. Ivanov, V. I., Zhurkin, V. B., Zavriev, S. K., Lysov, Yu. P., Minchenkova, L. E., Minyat, E. E., Frank-Kametskii, N. D. & Schyolkina, A. K. (1979) *Int. J. Quant. Chem.* **16,** 189–201.
28. Zhurkin, V. B., Lysov, Yu. P. & Ivanov, V. I. (1978) *Biopolymers* **17,** 377–418.
29. Sukhorukov, B. I., Gukowsky, I. Ya., Pekrov, A. I., Gukowskaya, N. M. (1980) *Int. J. Quant. Chem.* **17,** 339–359.
30. Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P. & Rich, A. (1976) *J. Mol. Biol.* **104,** 109–144.
31. Rosenberg, J. M., Seeman, N. C., Day, R. O. & Rich, A. (1976) *J. Mol. Biol.* **104,** 145–167.
32. Camerman, N., Fawcett, J. K. & Camerman, A. (1976) *J. Mol. Biol.* **107,** 601–621.
33. Klug, A., Jack, A., Viswamitra, M. A., Kennard, O., Shakked, Z. & Steitz, T. A. (1979) *J. Mol. Biol.* **131,** 669–680, and references therein.
34. Bloomfield, V. A., Crothers, D. M. & Tinoco, I., Jr. (1974) *Physical Chemistry of Nucleic Acids*, Harper & Row, New York, Chap. 5.
35. Shellman, J. A. (1974) *Biopolymers* **13,** 217–226.
36. Olson, W. K. (1980) *Nucleic Acids Geometry and Dynamics*, Pergamon Press, New York, pp. 383–398.
37. Record, M. T., Jr., Anderson, F. C. & Lohman, T. M. (1978) *Q. Rev. Biophys.* **2,** 103–178.
38. Manning, G. S. (1978) *Q. Rev. Biophys.* **2,** 179–246.
39. Clementi, E. & Corongiu, G. (1981) IBM Tech. Rep. POK-1.

Received December 16, 1980
Accepted May 22, 1981

(79)

Simulations of the Solvent Structure for Macromolecules. III. Determination of the Na^+ Counter ion Structure

ENRICO CLEMENTI and GIORGINA CORONGIU, *IBM DPPG, Poughkeepsie, New York 12602*

Synopsis

We report on a computer experiment in which, using Monte Carlo techniques, we considered a three-turn (30-base-pairs) B-DNA fragment as a solute and a set of 1200 water molecules and 60 sodium counterions (at a temperature of 300 K) as a solvent. From a statistical analysis of the Monte Carlo simulation (applied to the water molecules and counterions in the B-DNA field), we determined that the counterions themselves conform to two helical structures intertwined with the two strands. The structures of the water molecules solvating both counterion helices and the two B-DNA strands are fully analyzed and described in detail. A model for base-pair recognition based on the above findings is proposed. Aspects of the unwinding mechanism are discussed.

INTRODUCTION

In the first paper of this series,[1] we reported the statistically determined positions and orientations of 447 water molecules solvating a fragment of B-DNA composed of 12 base pairs (and of the corresponding phosphates and sugar units). In the second paper,[2] Na^+ counterions (one ion/phosphate group) were added to the solvent–solute system; it was assumed that the location of the Na^+ ions relative to B-DNA was the same as the one determined from quantum-mechanical computations of the interaction of one Na^+ ion and one sugar–phosphate fragment.[3] This assumption, however, was only an approximation, since an Na^+ ion in DNA interacts not only with one sugar–phosphate unit, but also with the entire B-DNA fragment. In this paper we drop this assumption and determine the statistical positions of each counterion in the total field: the rigid B-DNA fragment, the "free" water molecules, and the remaining Na^+ ions. This computer experiment makes use of Monte Carlo sampling techniques as proposed by Metropolis et al.[4] In this technique, the molecules constituting the ionic solvent (ions and water molecules) are randomly displaced (i.e., a "random walk") within a given volume containing the solute (B-DNA), which is assumed to be rigid. The simulated temperature in our experiment is 300 K.

EXPERIMENTAL METHOD

The selected B-DNA double-helix fragment is composed of 30 base pairs (three full B-DNA turns). Periodically, at each turn, the following sequence of 10 base pairs is selected for one strand: TA*, GC*, AT*, CG*, TA*, GC*, CG*, AT*, TA*, and CG*, where the asterisk denotes the h* strand. The coordinates of this B-DNA fragment have been discussed previously.[5] The three B-DNA turns (Fig. 1) are hereafter referred to as top, middle, and bottom turn, respectively.

Four hundred water molecules and 20 Na^+ counterions are placed in the middle section of the cylindrical volume, as shown in Fig. 1. For each water molecule we compute the interaction energy of the full three turns of the B-DNA fragment with the water molecules and Na^+ ions. In addition, since our B-DNA fragment has a periodically repeated sequence for each turn, each water molecule (or each ion) is associated with two "image" water molecules (or two "image" ions) obtained by a coordinate translation (along the Z axis) of the water molecule (or ion) in the central section. In Fig. 1, we show a water molecule in the central section of the cylindrical volume and its two "images" located symmetrically in the top and bottom sections. Equivalently, for each counterion in the central section, we compute the interaction of the atoms in the three DNA turns with all the water molecules and their corresponding water images, and with the other ions and their ion images. Thus, we simulate a system composed of the atoms in the three B-DNA turns, 1200 water molecules, and 60 counterions. The solvent particles (water molecules and sodium counterions) in the central section are randomly displaced[4]; the displacement applied to a solvent particle is also imposed on the "image" particles. Each random displacement (or "move") generates a new "configuration" for the solvent (or a new "step"

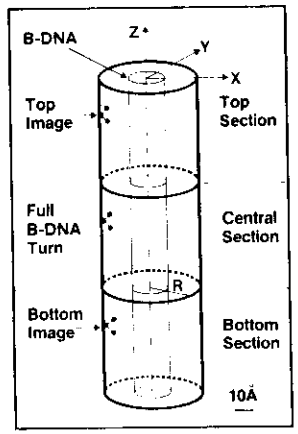


Fig. 1. Volume enclosing the B-DNA fragment: "image water" molecules and "image" ions.

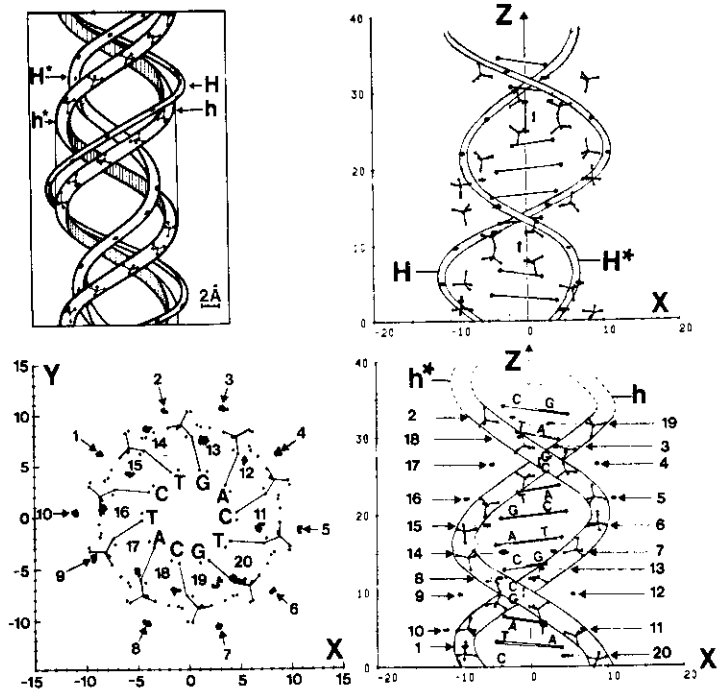


Fig. 2. (Bottom left) Projection of the statistical distribution of the 20 counterions in the X-Y plane. (Bottom right) Same distribution projected in the X-Z plane. (Top right) Same as bottom right but with connecting helices for the ions, rather than the phosphates. (Top left) Combination of top-right and bottom-right projections for more than one B-DNA turn.

in the random walk). An initial set of about 10^6 conformations was disregarded, however, in order to "erase" any memory of the initial configuration for the solvent. Thereafter, the Cartesian coordinates of the solvent particles (for each configuration) and the corresponding interaction energies (water with water, ions with DNA, and ions with ions and DNA) are stored on magnetic tape to be used later in a statistical analysis of the Monte Carlo data.

Analysis of the Monte Carlo Data

Two types of information can be obtained from the analysis of the Monte Carlo random walk, namely, *structural* and *energetic* information. In this paper, we report mainly on the structural data analysis. For each water molecule, and at each configuration, we compute the distances of the hydrogen and oxygen atoms to all the atoms of the B-DNA fragment and to

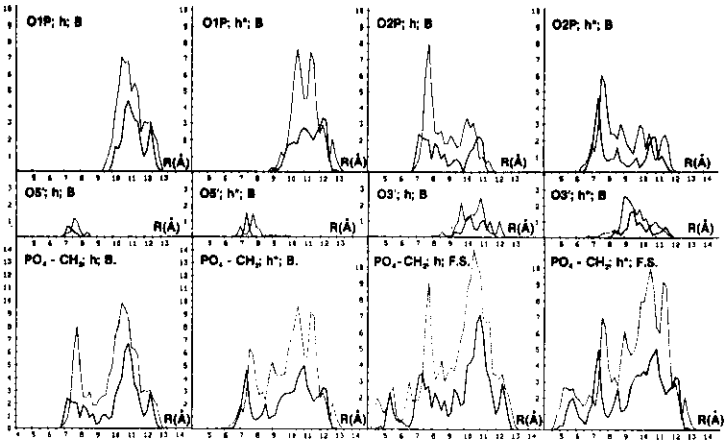


Fig. 3. (Top and central inserts) Distribution of the hydrogen and oxygen atoms of water molecules bound to the oxygen atoms in the PO_4^- groups. (Bottom inserts) equivalent distribution for the PO_4^--CH_2 group, either considering bound (first two inserts from the right) or first solvation shell water molecules (last two inserts from the right).

the counterions. The same type of computation is carried out for the ions; the distances are then analyzed. A water molecule is defined as bound to an atom of DNA (or to a counterion) if its distance and orientation satisfies

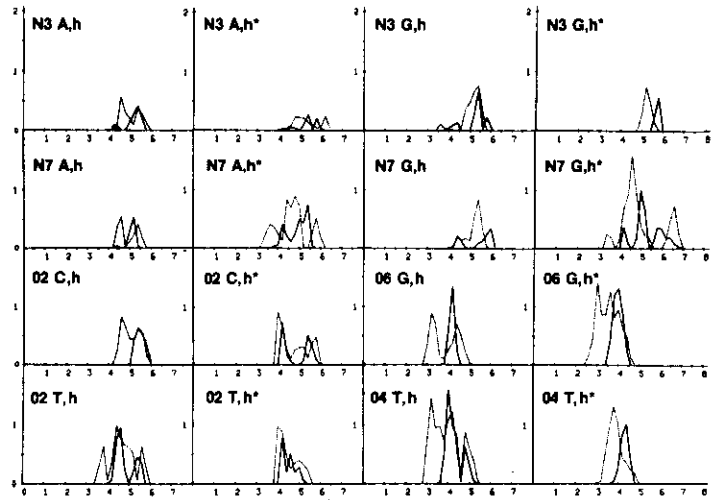


Fig. 4. Distribution of hydrogen and oxygen atoms for water molecules bound at the nitrogen (N3 and N7) or oxygen (02, 04, and 06) atoms of bases forming either the h or the h* strand in B-DNA.

the predetermined geometrical criteria previously discussed.[1,2] A water molecule is defined as being in the first solvation shell of an atom in the DNA fragment (or of a counterion) if the distance criteria (see above), but not the orientation criteria, are satisfied. Since the atoms of the DNA fragment can easily be distinguished as belonging to either the first or second B-DNA strand, we can use this analysis to differentiate the statistical probability that a water molecule will interact with one or the other strand.

A water molecule can assume a position such that it is bound to more than one atom or exists in the first solvation shell for more than one atom. Physically, this is the case for water molecules bridging two atoms. For this reason, and also in order to obtain additional statistical data, we performed the above analysis not only for water molecules interacting with DNA atoms, but also for groups of atoms (i.e., the four oxygen atoms of the phosphate group or all the atoms of a given base). We recall[1,2] that the number of water molecules bound to a group of atoms is generally smaller and, at most, equal to the sum of the number of the water molecules bound to the atoms constituting the "group of atoms."

When a water molecule is neither bound nor in the first solvation shell of either an atom in DNA or a sodium ion, then the molecule is defined as "groove water". Making use of the geometrical characterization of a groove,[2] we can further differentiate between major and minor grooves.

It is noted that the computational time required to perform the above analysis amounts to a significant fraction of the total Monte Carlo simulation, and its programming is more extensive and complex than the one required for the Monte Carlo random-walk simulation. In the following, we report the statistical data obtained from about 10^6 conformation, which were generated after disregarding the 10^6 needed to erase the memory of the initial configuration and to equilibrate the solvent system; the initial position for the ions and the water molecules was the final configuration previously obtained.[2]

Determination of Counterion Structure

In a recent, preliminary communication,[6] we reported that the sodium counterions in a solution with B-DNA form a pattern corresponding to two helices intertwining with the two B-DNA strands. The overall methodological approach is described at length elsewhere,[3] as well as having been previously summarized.[7]

The most direct way to analyze the counterion positions is to provide the projections in the X-Y and X-Z planes (see Fig. 1 for the choice of axis) of the statistical distributions of the counterions obtained from the Monte Carlo data. The following technique has been adopted: the entire cylindrical volume (Fig. 1) was subdivided into small cubical cells (0.2 Å/side). A counter at each cell was activated to measure how many times a counterion falls within a cell during the Monte Carlo walk. The statistical

distribution is graphically visualized by reporting for each cell a number of points proportional to the number of times an ion is present in the cell. The projections of the probability distributions of the counterions at one B-DNA turn are reported in Fig. 2. In the bottom-left insert we provide the X-Y projection of the ion distribution map (spotlike patterns), with an index of 1–10 for the 10 "spots" external to the phosphates and an index of 11–20 for the 10 "spots" internal to the phosphates. A counterion corresponds to each "spot"; the size of the "spot" provides a measure of counterion mobility. The phosphates of one strand are indicated by drawing the bonds between the oxygen atoms and the corresponding phosphorus atom and by explicitly indicating the corresponding bases, A, G, C, and T. The atoms of the phosphates in the second strand are indicated simply with dots at the nuclear positions, and for the corresponding bases, we report only a dot for the nitrogen atom position (the one connecting the base with the sugar). The projections of the 20 counterion probabilties form two nearly regular circular patterns. In the bottom-right insert of Fig. 2, we present the same probability density distributions—this time projected into the X-Z plane. In this insert, the phosphate groups of the two strands are enclosed in helical envelopes, and the base pairs are identified by reporting the base-pair molecular plane. In the top-right insert, we repeat the X-Y projection—this time connecting to an envelope those counterion probability distributions that are nearest neighbors. To simplify the diagram we have not enveloped the phosphate groups in two helices as done in the bottom-right insert. The pattern emerging from these three depictions is repeated in the top left insert. There we draw the two helical envelopes for the phosphates of the two strands and the two helical envelopes for the sodium counterions so that one penetrates the major groove and the other is outside the minor groove. The counterion helices are designated by the letters H and H*. It is evident from the density projections that the H helix is external to the cylindrical volume determined by the phosphate groups, whereas the H* helix is internal to it. The cross section of the imaginary "cable" enveloping the ions of H* is larger than the corresponding cross section of the "cable" enveloping the ions of H. The physical reason is rather obvious: the counterions in H* are strongly afected by the base pairs. Therefore, the exact position of an ion in H* is base-pair-sequence dependent, and the "irregularities" of distributions 11–20 in the bottom-left insert of Fig. 2 are the effect of this base-pair sequence dependence.

The new structure, H and H*, obtained in our computer experiment is physically very satisfying since it simultaneously optimizes several basic energy requirements: it keeps the counterions as far as possible from each other but at the same time satisfies the very strong attraction between the Na^+ ions and the free oxygen atoms of the phosphates and the strong attraction between the Na^+ ions and the bound oxygen atoms in the phosphate groups and the attraction to the base-pairs; finally, it allows both the phosphate groups and the counterions to be solvated by the water molecules, thus making use of the solvation energy to stabilize the entire system.

Biologically, this structure is interesting since (1) it provides for a base-pair-recognition mechanism at long range due to the base-pair-sequence dependent H* structure; (2) it can easily allow for the exchange of one or more sodium counterions with different counterions (from Na^+ to K^+ or Mg^{2+} or Ca^{2+}, etc.); (3) it can account for rapid structural and conformational reorganization processes typical of ionic solution with macromolecules as solute; and (4) because of the ionic mobility and very strong interactions, it can act as either an important sink (or source) of energy.

Determination of the Water–Structure Solvating DNA

Statistically, the most probable distribution of the water molecules (at 300 K temperature), either bound or in the first solvation shell, for an atom (or groups of atoms) in B-DNA is analyzed in Figs. 3–7. In these figures, the probability distribution for the hydrogen atoms is given as a dotted line and the one for the oxygen atoms is given as a solid line. On one axis (ordinates) we report the number of hydrogen or oxygen atoms as a function of R (given in the abscissa, in Å); we recall that R is the distance of the hydrogen (or oxygen) atoms from the Z axis (see Fig. 1). The notations B and FS differentiate between bound and first solvation shell water molecules; the notations h and h* differentiate the two strands. In Table I we summarize the R values for some of the atoms of B-DNA (discussed in detail in the following).

In Fig. 3 we report the analysis for water molecules bound to the free oxygen atoms (O1P and O2P) to the bound oxygen atoms (O5′ and O3′) in the two strands (h and h*). These analyses of water molecules bound at atomic sites are complemented with the bound water distributions and the first solvation shells distributions at the PO_4^--CH_2 group.

We learn that at the O1P sites the water oxygen atoms in the h strand have a maximum at about 11 Å, whereas the maximum is at 12 Å in the h* strand. The water hydrogen atoms distributions (dotted lines) in h differ

TABLE I
R Distance (in Å) for Selected Atoms in B-DNA

Atom	R	Atom	R
O1P	10.2	H of NH_2 in G	4.0
O2P	8.8	H of NH_2 in G	3.0
O3′	8.8	N3 in A	3.2
O5′	7.7	N7 in A	4.0
P	8.9	N3 in G	3.3
O4′	6.2	N7 in G	3.9
H of NH_2 in A	2.8	O2 in C	3.7
H of NH_2 in A	1.8	O6 in G	1.7
H of NH_2 in C	2.1	O2 in T	3.6
H of NH_2 in C	3.7	O4 in T	2.8

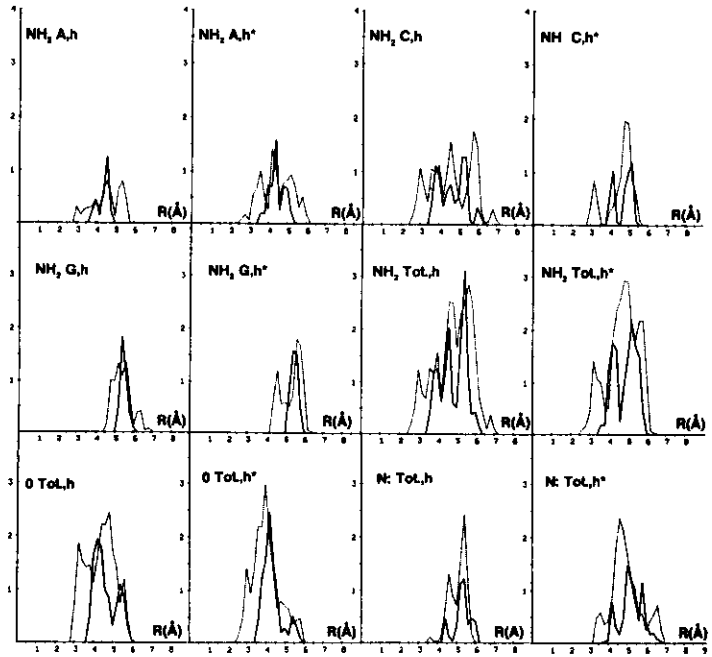


Fig. 5. (Top and central inserts) Distribution of hydrogen and oxygen atoms for water molecules bound at the NH_2 groups of the bases. (Bottom inserts) Equivalent quantities but for an average oxygen (either 02, 04, or 06) atom and an average nitrogen (either N3 or N7) atom in the bases at the two antiparallel strands h and h*.

from those in h*, which indicate different orientational arrangements at the two strands. The same holds for the water molecules at O2P. There is little, if any, water at the O5′, as previously noted,[1,2] and not much water at O3′. This information is iterated by providing the distribution for the water molecules bound at the PO_4^--CH_2 site. As previously discussed,[2] the first solvation shell can extend much further than the bound water distribution, as is clearly seen in Fig. 3. The integral of the distribution values of the hydrogen and oxygen atoms as a function of R provides the number of water molecules bound to the atoms or groups of atoms (Fig. 3). These are given in Table II for both bound and first hydration shell water molecules. In the following, we generally do not comment on features that can be obtained by inspection of the figures. We feel that ir, Raman, nmr, neutron, and x-ray experiments on DNA in solution will be more easily interpreted given the availability of these data.

In Fig. 4 we report the analysis for the water molecules bound to the lone-pair nitrogen (N3 and N7) and oxygen (O2, O4, and O6) atoms of the bases. The analysis of the water molecules bound to atoms belonging to

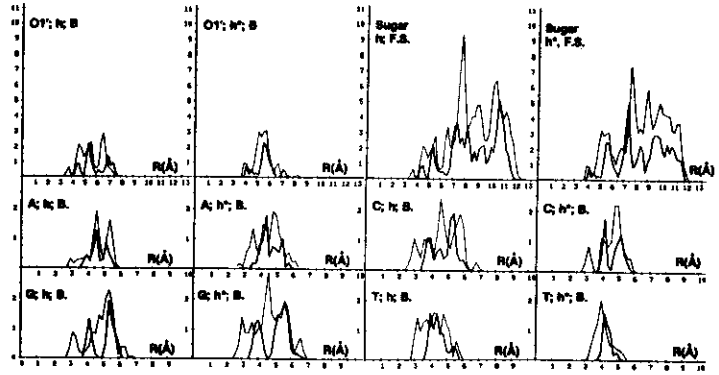


Fig. 6. (Top inserts) Water characterization at the sugar units in the two strands: bound water at 01′ atom and first solvation shell water in the sugar units. (Middle and bottom inserts) Distribution for water molecules bound to the bases.

the bases is extended by reporting (Fig. 5) the distributions for the water molecules bound to specific NH_2 groups. The average distributions for water molecules bound to NH_2 or oxygen or nitrogen atoms at the bases is obtained by considering all the bases at the two strands.

In Fig. 6 we compare the distributions of water molecules bound at the O1′ (of sugar) or present in the first solvation shell of the entire sugar unit. This figure also shows the distribution of the water bound to the bases (in each strand).

Fig. 7 we complete the analysis of the bases by reporting the average distribution of the water molecules bound to the A-T and G-C base pairs, the average distribution for the base pairs in the h and h* strands, and the first hydration shell distribution at the bases (average values for both strands).

These very detailed but clear graphical presentations of the solvation in B-DNA are complemented by the data in Table II.

Determination of the Water Structure Solvating the Counterions

In Fig. 8 we report the distribution of the water molecules bound to the 10 counterions in the H helix (ions 1–10) and to the 10 counterions in the H* helix (ions 11–20). In this figure the counterion is placed at the origin of the axis. The orientation of the water molecule (oxygen nearer, hydrogen farther away) is very typical of a sodium ion in solution.[8–10] Each counterion is solvated; this finding is expected to the valid also for K^+ counterions, and most likely for Li^+ counterions as well.

In Fig. 9 we report the distribution for the water molecules in the two strands, either bound or in the first hydration shell. From these data we see clearly that the two strands have a different hydration pattern, as is also reported in Table II. This very important finding remains hidden by only

TABLE II
Number of Bound Water Molecules and Their Interaction Energies with Atoms or Groups of Atoms of B-DNA and the Na^+ ions

Atoms or Groups	Number of Bound Waters		Average Energy (kcal/mol)	
	h	h*	h	h*
O1P	3.24	3.36	−37.97	−34.57
O2P	2.92	3.05	−38.33	−41.11
O5′	0.18	0.23	−37.76	−43.52
O3′	0.89	1.12	−37.51	−34.24
O1′	1.19	1.13	−39.00	−37.81
N3 in A	0.49	0.26	−39.07	−34.97
N7 in A	0.46	0.94	−37.97	−37.73
N3 in G	0.72	0.29	−36.70	−39.74
N7 in G	0.60	1.20	−33.03	−41.94
O2 in C	0.67	0.97	−39.91	−37.33
O6 in G	1.08	1.34	−35.91	−40.53
O2 in T	1.00	1.00	−41.08	−38.94
O4 in T	1.62	1.26	−36.97	−29.29
Na^+	4.00	4.40	−39.72	
PO_3^--CH_2	6.43	7.18	−38.04	−37.09
NH_2 in A	1.49	1.77	−39.79	−37.65
NH_2 in C	2.76	2.31	−38.24	−39.18
NH_2 in G	1.94	1.46	−36.23	−39.93
NH_2 in A, G, C	2.13	1.70	−37.81	−39.04
N in A, G	0.57	0.67	−36.40	−38.85
O in C, G, T	1.00	1.06	−38.61	−38.35
A	2.42	2.55	−40.03	−37.46
C	3.43	3.29	−38.58	−38.43
G	3.84	4.15	−35.77	−40.68
T	2.62	2.26	−38.51	−37.38
Grooves	131.82		−30.41	

reporting the distributions for the total system of 400 water molecules solvating the 10 base-pair turns in B-DNA or by considering the water molecules in either the first solvation shell or those bound to B-DNA. In Fig. 10 we present the distribution of water molecules bound to B-DNA and the ions in the first hydration shell and in the grooves.

In Fig. 11 we report the average energy for a water molecule in volume R and $R + dR$; the energy is decomposed as water–water, water–DNA, and water–ions. Notice how this energy is nearly constant from small to large values of R.

CONCLUSIONS

The most important conclusion from our computational experiment is that the counterions and the solvation water molecules form two different patterns at the two DNA strands. We designate the global system composed by the h strand, the counterions in the H helix, and the water molecules bound to h and H, as the S "superstrand." The equivalent global system for h* and H* and the solvating water molecules is designated as

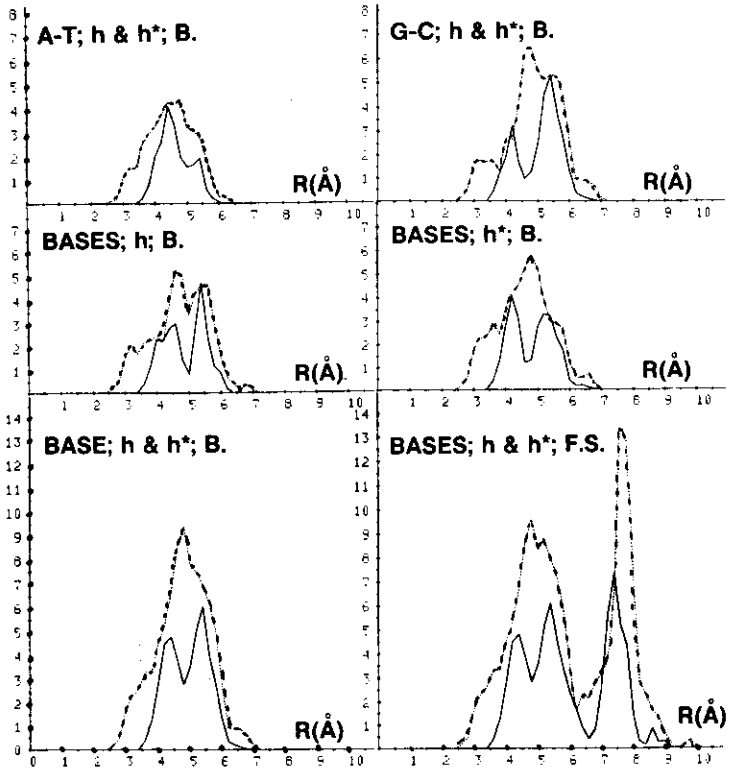


Fig. 7. (Top inserts) Distribution of water molecules at the base pairs. (Middle inserts) Average distribution of water molecules bound in the two strands. (Boottom inserts) Average distribution for both strands of water molecules bound (left) and in the first solvation shell (right) of the bases.

S*. As known, the two strands h and h* differ only because they are antiparallel. This difference, however, is enhanced by the counterion distribution and by the water molecules. Any biological process dealing with DNA in water solution involves the S and S* superstrands and not just with the h and h* strands. In addition to a given DNA conformation, there is a correspondingly wide spectrum of S and S* conformations. Indeed, the structure of H and H* is dependent on the counterion charge, the ionic radius, and the counterion concentration for a given temperature.

In the following we consider a few immediate implications of the above findings. We present a "base-pair sequence" recognition mechanism. Let us consider for example, a molecule of glycine zwitterion, approaching DNA in water solution but still relatively far away from DNA, such that the direct interaction glycine–base pairs can be assumed to be small. We assume that

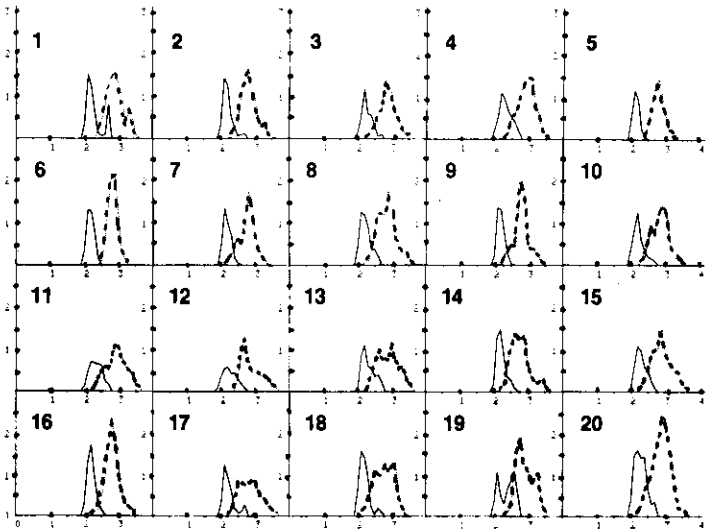


Fig. 8. Hydrogen and oxygen atom distributions for water molecules bound to the counterions in the H helix (1–10) and in the H* helix (11–20).

the C^{α} of glycine is at 17 Å on the x axis (with $y = 0$; see Fig. 1) and is optimally oriented relative to our B-DNA fragment, the 1200 water molecules, and the 60 sodium counterions. Further, we assume that glycine can be translated along the z axis, reoptimizing its orientation at each value of Z.

The interaction of the counterions with glycine is rather large (long-range interaction of an ionic nature), and the interaction of the H* counterions with the base pair has been shown previously to be large and base-pair sequence dependent. Therefore, glycine will recognize the base-pair sequence, via the counterions. The proposed recognition mechanism is of the relay type: base pairs to counterions, counterions to glycine. It is noted that the water solvent molecules amplify the recognition ("signal amplification") because of the shell structure and other hydrogen-bonded structures (see, for example, some recent comparative studies of methane[11] and methanol[12] in water solutions). Among the applications of this proposed mechanism, we mention (1) recognition of a sequence perturbed by cancer or an anticancer intercalating molecule and (2) recognition of amino acids by RNA in protein syntheses.[6]

Another implication of our findings concerns with unwinding mechanism in the double helix. As is known, a double-helix structure has a critical temperature and a critical ionic concentration outside which the two helices snap apart extremely rapidly. With a 20 K increase in the temperature (in our simulated system), we obtained a different counterion pattern, with the counterions in the H* helix closer to the base pairs than at 300 K. We

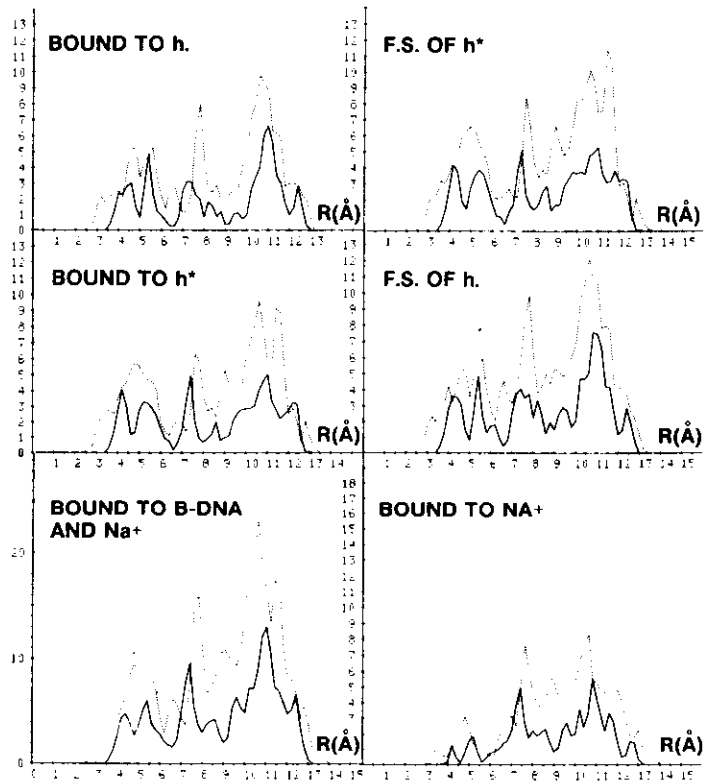


Fig. 9. (Top and central inserts) Average hydrogen and oxygen atom distributions for water molecules in the h and h* B-DNA strand either bound or in the first hydration shells. (Bottom inserts) Water molecules bound to B-DNA and to the counterions (left) or only to the counterions (right).

recall that the interaction of an Na^+ ion with the base pairs is not only strong when Na^+ is at the perimeter of the basepair (in the plane containing the base-pair skeleton), but also when Na^+ is above the base pair. In this position, the attraction of the "base pair to Na^+" is opposed by the hydrogen atoms forming the base-pair hydrogen bonds. An increase in the system's thermal motion (due to temperature) can bring about a separation between two successive base-pair and/or hydrogen bond breakage within a base pair. In either case, a sodium ion can approach the bases even further and oppose the restoration of the original DNA configuration. Thus, we conclude that the disruption of the double helix following a temperature increase is due not only to the increased amplitude of the DNA vibrational modes, but also to an increasingly strong anharmonicity of such modes caused by the approach of counterions to the base pairs.

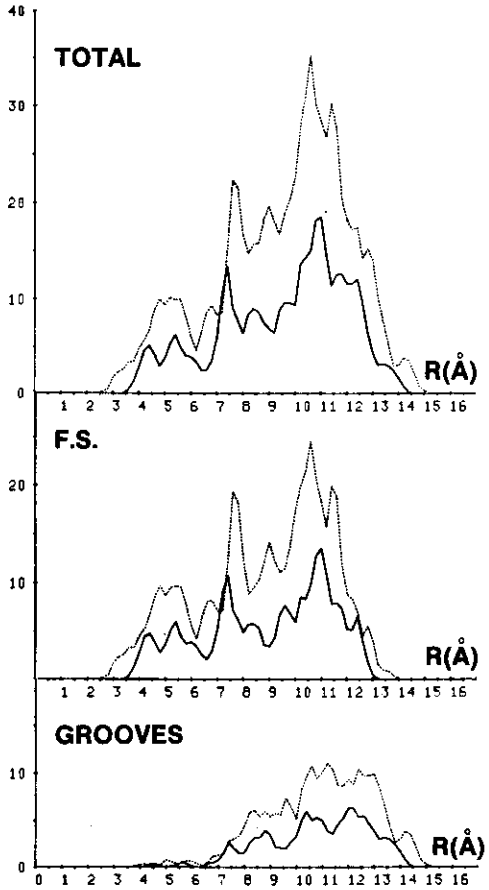


Fig. 10. Statistical hydrogen and oxygen atoms total distribution for 400 water molecules solvating one B-DNA turn and 20 Na^+ counterions (top) in B-DNA first solvation shell (middle) and in the grooves (bottom).

The disruption of the double-helix structure following progressive removal of counterions at a constant temperature is easily understood in terms of the strong stabilization brought about by the ions to the "water and DNA" system.[2] We note that the total energy of the system reported in this work is more stable by about 20 kJ/mol than the system analyzed in Ref. 2.

On the basis of our energy data, we estimate that the DNA and counterion field extends up to about $R = 25$ Å. Therefore, we expect that x-ray crystal studies from single crystals should show evidence of DNA-to-DNA

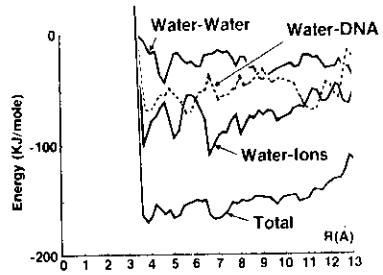


Fig. 11. Average interaction energy for water with the remaining water molecules (water-water), with the B-DNA fragment (water-DNA), and with the ions (water-ions) in the total interaction as a function of R.

perturbation. As a consequence, the counterion structure in a single crystal is expected to differ from the counterion structure of DNA in solution.

A study is in progress to determine the counterion structures of Li^+, K^+, Mg^{2+}, and Ca^{2+}.

The research of G. Corongiu was partially sponsored by the National Foundation for Cancer Research.

References

1. Corongiu, G. & Clementi, E. (1981) *Biopolymers* **20**, 551–571.
2. Corongiu, G. & Clementi, E. (1981) *Biopolymers* **20**, 2427–2483.
3. Clementi, E. (1980) *Computational Aspects for Large Chemical Systems*, & *Lecture Notes in Chemistry*, Vol. 19, Springer-Verlag, New York.
4. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. (1953) *J. Chem. Phys.* **21**, 1087–1092.
5. Clementi, E. & Corongiu, G. (1980) *J. Chem. Phys.* **72**, 2979–3992.
6. Clementi, E. & Corongiu, G. (1981) *Biomolecular Stereodynamics*, Sarma, R. H., Ed., Adenine Press, New York, pp. 209–259.
7. Clementi, E. (1981) *IBM J. Res. Dev.* **25**, 315–325.
8. Clementi, E. & Barsotti, R. (1976) *Theor. Chim. Acta* **43**, 101–120.
9. Clementi, E. & Barsotti, R. (1978) *Chem. Phys. Lett.* **59**, 21–25.
10. Clementi, E. (1976) *Determination of Liquid Water Structure, Coordination Numbers for Ions and Solvation for Biological Molecules, Lecture Notes in Chemistry*, Vol. 2, Springer-Verlag, New York.
11. Bolis, G. & Clementi, E. (1981) *Chem. Phys. Lett.* **82**, 147–152.
12. Corongiu, G. & Clementi, E. (1982) *Chem. Phys. Lett.* (in press).

Received April 30, 1981
Accepted September 4, 1981

Solvation of DNA at 300 K: Counter-ion Structure, Base-pair Sequence Recognition and Conformational Transitions A Computer Experiment

Enrico Clementi and Giorgina Corongiu[61]
International Business Machine Corporation
Poughkeepsie, N. Y. 12602

Introduction

Previously we have reported on the interaction between one water molecule and the bases[1] and base-pairs[2] of the nucleic acids, A-DNA single helix[3], B-DNA single[4] and double helices[4,5]. In addition, *Monte Carlo simulations* (at 300 K) have been presented for a cluster of water molecules enclosing the bases and the base-pairs[2], or a limited region around A-DNA single helix[3] and B-DNA single helix.[4-6] *These studies represent preliminary steps.* We extend our previous effort by considering, via simulations, not only a much larger number of water molecules than previously studied, but also the effect of counter-ions, initially the Na^+ ion. We shall discuss B-DNA interacting with either a few or many water molecules per nucleotide unit, presenting our findings as a *set of three successive approximations.* In this computer experiment the simulated temperature is 300 K. *At first* we considered B-DNA and water molecules; *secondly,* we consider the same system, but to each phosphate unit we add a sodium counter-ion placed at a *pre-determined* position. *Finally* we shall consider a larger B-DNA fragment and the water molecules and the ions will be let *free* to assume the statistically optimal positions and orientations for the temperature analyzed. This third step brings about a newly determined structure for the counter-ions. At each approximation we consider the full range of relative humidities.

This study is applied to *conformational transitions.* On the basis of new findings (counter-ion structure) we present a base-pair sequence *recognition mechanism* and a quantitative simulation.

Graphical Representation of the Probability Density

Some of the analysis reported below is not carried out using probability density maps, as done in the past[2-9] since these are somewhat difficult to read. The probability maps have been replaced by an algorithm described in the following four steps: 1) after computation of the probability density maps, the probability density maxima for the oxygen atoms are located, 2) for the hydrogen atoms, the proba-

bility maxima are determined subject to the constraint of being located on a sphere of radius equal to one half of the O-H internuclear separation in H_2O, 3) a sphere of radius 0.5 A is centered at the oxygen and at the two hydrogen atoms probability maxima, and 4) the conformations from the Monte Carlo simulation are scanned to determine how many times a water molecule fell into the volume defined in Ref.3. We note that the assumption of a sphere around the oxygen and hydrogen atoms, implies an isotropic probability distribution; as pointed out previously[2-4] the probability distribution is often anisotropic, especially at room temperature. On the other hand, the advantage of the new method is that it allows a graphical representation of immediate understanding. The technique described here brings 60% to 80% of the full set of the simulated conformations into the three spherical volumes associated to each water molecule. This representation is hereafter referred to as *"the average configuration"*.

The water structure at high humidity is analyzed in greater detail, since it corresponds to a situation expected to be rather near to the one for Na^+-B-DNA in a physiological solution; low humidity is of interest mainly for comparison with physico-chemical (rather than biological) experiments, for example on DNA fibers. The two strands and the constituent chemical groups are differentiated in this paper by presence or absence of an asterisk; thus, for example, P(n) and P*(n), G(n) and G(n)* refer to the n-th phosphate group, or to n-th base (guanine) of either the h or the h* helix, namely the two strands forming the double helix (h and h* are short notation for the 5′-3′ and 3′-5′ strands, respectively). Water molecules are labelled with an index and, in general, a water molecule with an index m has its oxygen atom *above* (projection on the Z axis) any water molecule with an index value larger than m.

We recall that in the first solvation shell one finds water molecules *"bound"* to sites (hydrophilic sites), and water molecules *near* the hydrophobic sites, generally acting as bridges between hydrophilic sites. Thus the *"first solvation shell"* distribution contains the *"bound"* distribution and additional water molecules. The difference between the *"total"* and *"the first solvation shell"* distributions provides the *"grooves"* distribution. It is clearly a matter of taste to speculate how much a "groove distribution" coincides with "second and third solvation shells distributions". We have opted for the "groove" terminology, since it allows further differentiation between major and minor grooves, a distinction of relevance in discussing nucleic acids. Lewin[10] subdivides the major groove's water molecules into three belts (upper, middle and lower); our subdivision is about equivalent, since the water molecules of the upper and lower belts correspond essentially to our *first solvation shell's* water molecules.

We note that since a water molecule experience the entire field of the DNA fragment, there is necessarily an element of ambiguity in deciding if a given water molecule "solvates" only a given site, unless an exact definition is provided. Our selection criteria to assign a water molecule as "bound" to a given atom "a" of the B-DNA fragment are: 1) the oxygen atom of water and the atom "a" internuclear

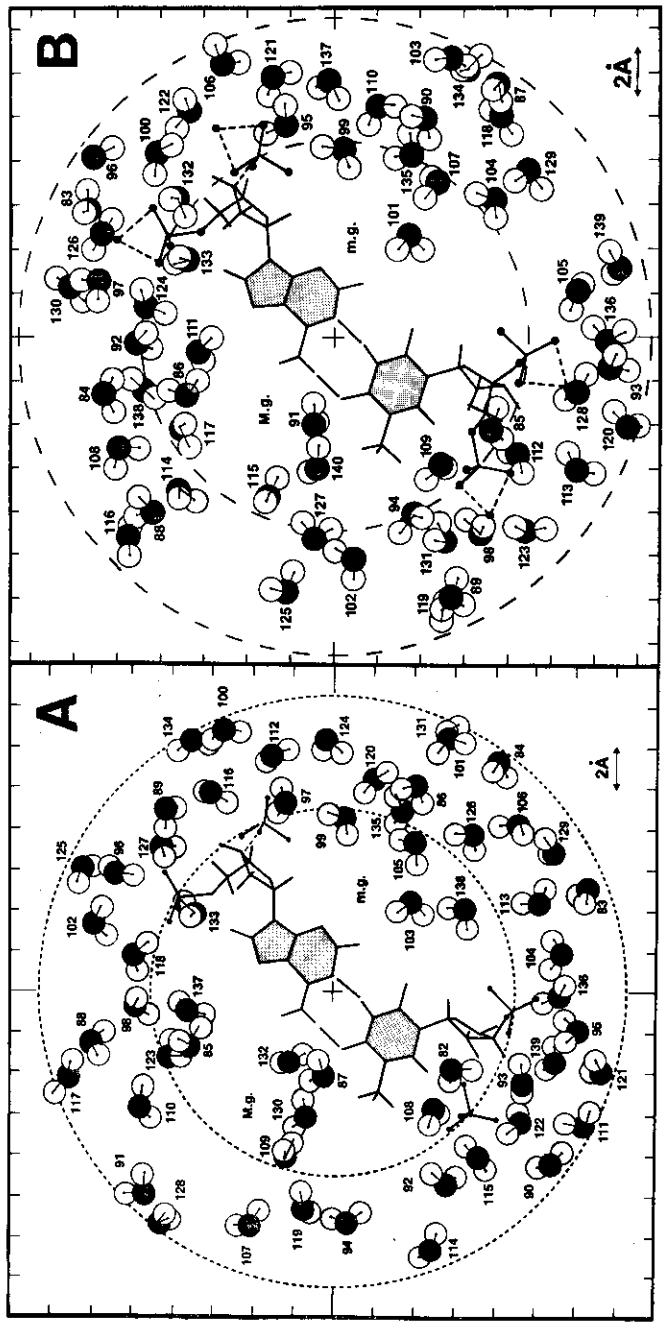


Figure 4. Water molecules contained in a disk of 4 A thickness and solvating B-DNA (insert A) and Na^+-B-DNA (insert B); the radii of the two circumferences are 14 A and 8 A, respectively.

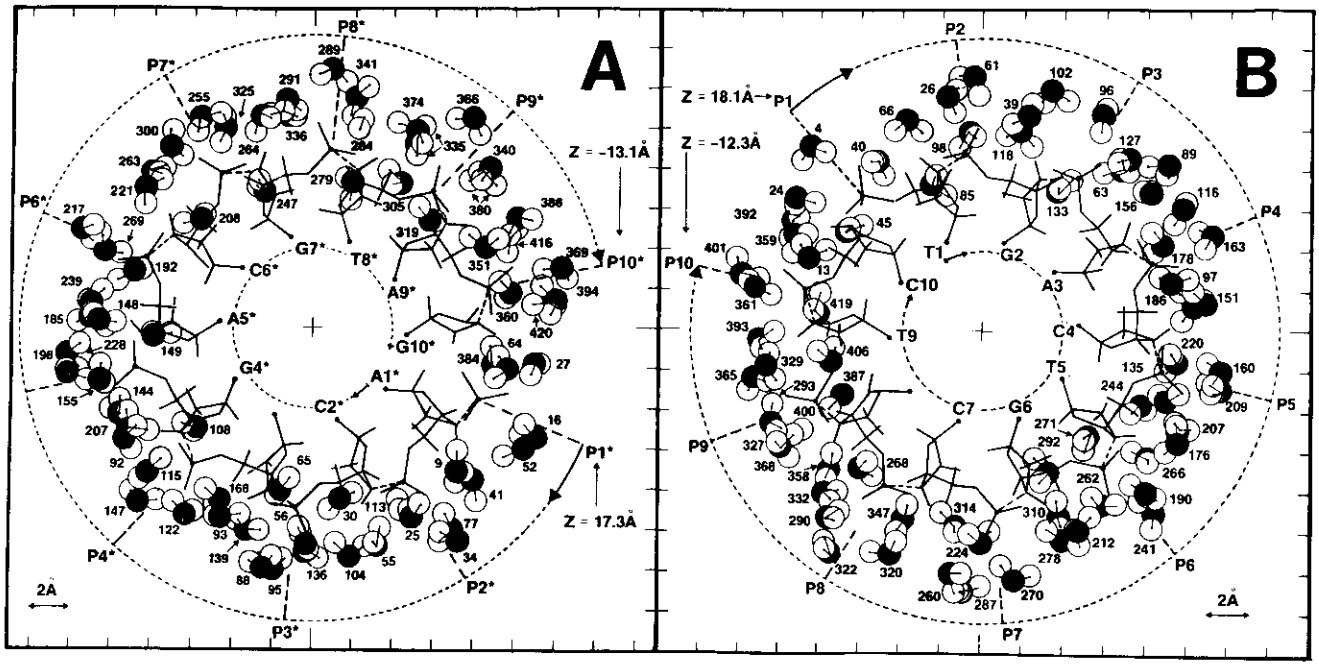


Figure 3. Water molecules solvating the h helix in (insert A) and the h* helix (insert B) in B-DNA.

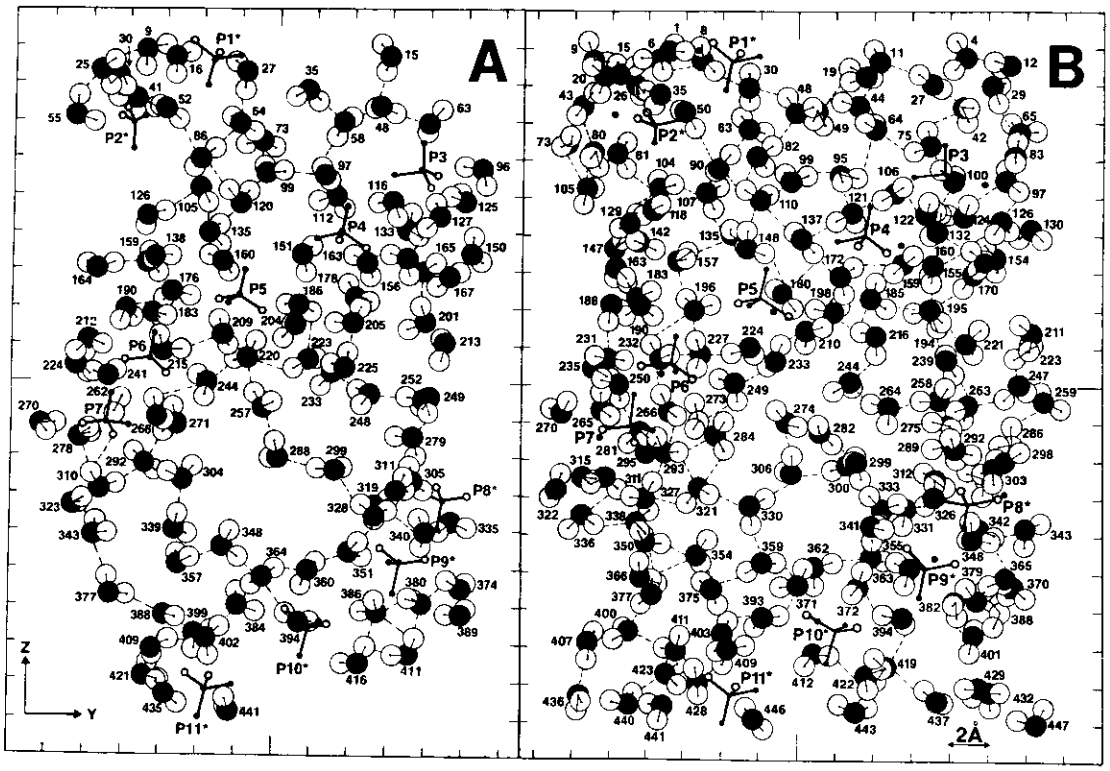


Figure 6. Network of water molecules in B-DNA and Na^+-B-DNA (see text).

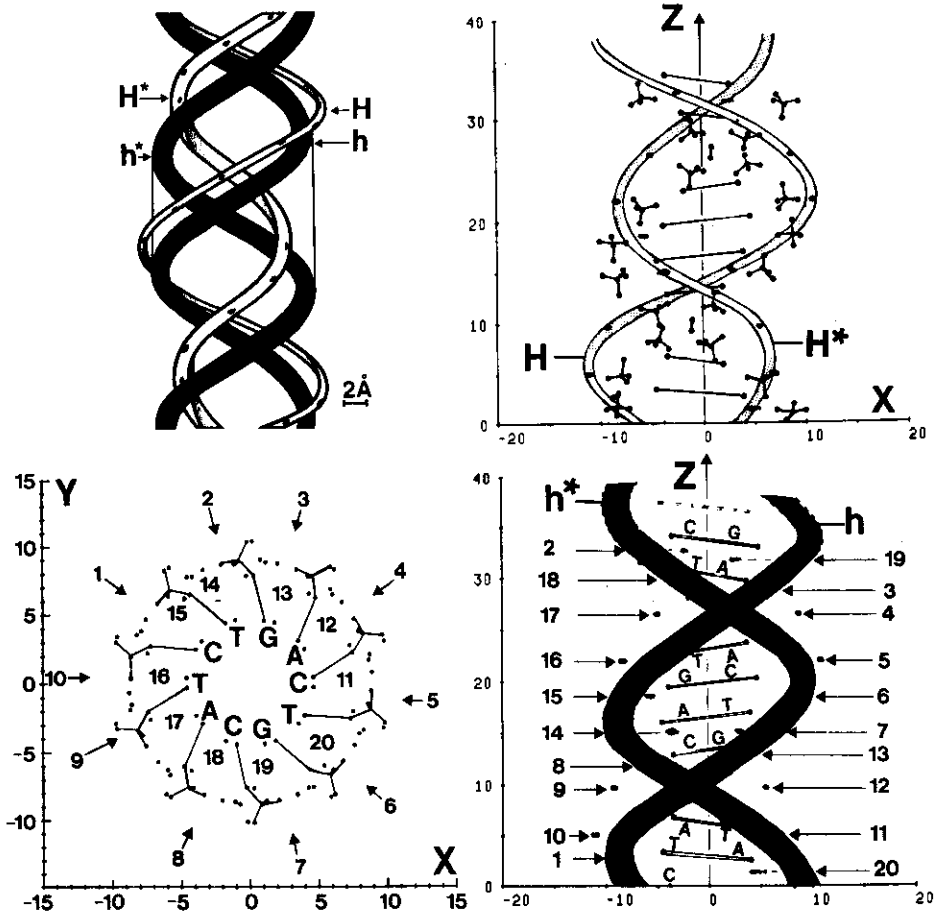


Figure 16. *Bottom-left:* projection of the statistical distribution of the twenty counter-ions in the X-Y plane. *Bottom-right:* same distribution projected in the X-Z plane. *Top-right:* same as bottom-right, but with connecting helices for the ions, rather than the phosphates. *Top-left:* combination of top-right and bottom-right for more than one B-DNA turn.

separation, R(a, O), must be smaller or equal to a threshold value T(a, O) and 2) one hydrogen atom of a given water molecule and the atom "a" internuclear separation, R(a, H) must be compared to a threshold value T(a, H) to ensure the proper orientation of that water molecule relative to the atom "a". In this way we can distinguish the case "a"—H-O from the case "a"—O-H (for example N: (in a base) from H of NH_2).

In Fig.1 we report the projection onto the X-Y plane of the third (A3-T3*) and fourth (C4-G4*) base-pair of our fragment (and the corresponding sugar-phosphate groups) in order to clarify our definitions of "*bound* water", "*first* solvation shell water" and "*groove* water." The thick solid line, *composed* by a family of circumpherences of radius T(a, O) and *enclosing* either A3-T3* or C4-G4*, represents regions of "bound" water molecules. A number of atoms "a" have been identified in the figure with a larger dot to represent the nuclear position and the radius origin. Some of the spheres are inside the thick solid line, like O3′ in A3-T3* and O2P, O5′, N3, N5 in C4-G4*. One can notice at a glance the strong overlap of these spheres, and this fact constitutes the main reason for the "*non-additivity*" namely, the total number of water molecules found to be *bound* to the DNA fragment is smaller than the sum of the numbers of water molecules found to be *bound* to the atoms composing the DNA fragment. However the "orientation" criterion imposed on the hydrogen atoms (threshold T(a, H)) reduces considerably the above inconvenience. The shaded areas at the periphery of A3-T3* and C4-G4* refer to water molecules in the *first solvation shell*, but not bound to hydrophilic sites (see for example the area for CH_3 of thymine in A3-T3*). Water molecules in this volume might be counted more than once in our analyses, once as being in the first solvation shell (for example of CH_3 for T3*) and once as being bound to a given nearby site. Finally, in C4-G4* we have presented two PO_3^--Na^+ groups near to guanine, G4*. Notice the overlap of O1P (at one phosphate group) with Na^+ (on the second group), or the overlap of O2P at one group with O5′ at the second group. Thus we stress once more that a water molecule *should not be* physically associated to a *single* site; its location and orientation are due to the *entire* field of the fragment. An assignment is, however, essential in order to interpret experimental data, in particular infrared, Raman and scattering data; and also it is very important to describe and understand the solvent structure.

First Approximation: B-DNA and Water Molecules

The B-DNA double helix fragment we consider has been previously discussed (see Ref.2), and consist of twelve base-pairs (namely, two more base-pairs than needed to reproduce a full B-DNA double helix turn) with the corresponding sugar and PO_4-CH_2 units. The following sequence of base-pairs has been selected: AT*, TA*, GC*, AT*, CG*, TA*, GC*, CG*, AT*, TA*, CG*, GC*. The B-DNA double helix fragment is enclosed into a cylinder with its axis co-axial to the B-DNA long axis (z axis). The cylinder height is 36.0 A with a base radius of 14.5 A. The two base-pairs and the corresponding sugar units have been added in order to improve the interaction field descriptions at the bottom and top ends of the B-DNA frag-

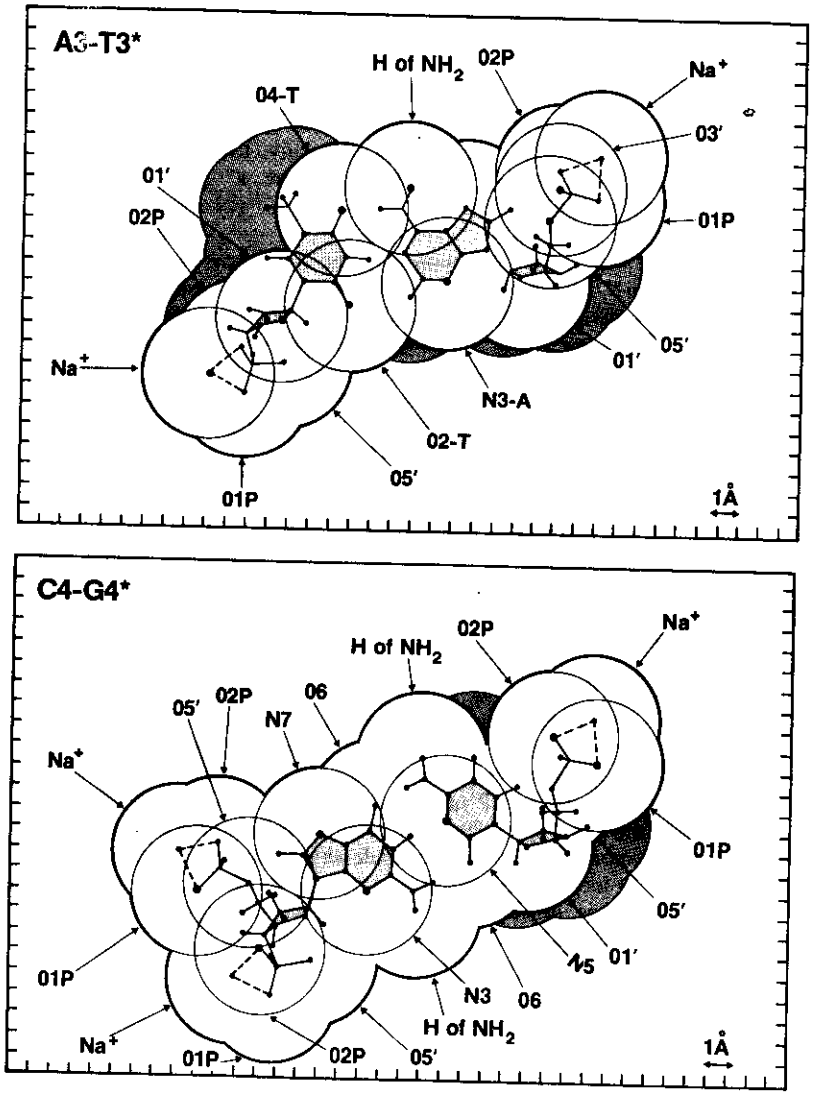


Figure 1. First solvation shell decomposition. Example for two subunits of the Na^+-B-DNA fragment.

ment. In the Monte Carlo simulation below reported 447 water molecules have been placed into the cylinder. The equilibration process was carried out for $2x10^6$ conformations; the statistical data below analyzed are obtained from additional

$2x10^6$ Monte Carlo "moves" (these computations have been carried out on either an IBM-370/3033 or a IBM-370/3081 computer).

Structure of First Solvation Shell

Let us start with a gross analysis of the computed data. In Figure 2 (insert A) we report the distribution of the internuclear distances from an atom (either O or H) of water to the nearest atom of the B-DNA fragment. The internuclear distance is designated as R(i-a), where "i" refers either to a hydrogen or the oxygen atom in water and "a" refers to the atom of B-DNA nearest to "i". We indicate as N (i) the number of atoms of type "i" located (from analyses of the Monte Carlo data) at a R(i-a) distance from "a". The R(O-"a") internuclear distance distribution (see the solid line of Figure 2) shows a very distinct peak with a maximum at 2.6 A (first region), a second peak with a maximum at 3.4 A (second region) and a set of peaks that approximately can be associated to an asymmetrical gaussian distribution with a maximum at about 4.5 A (third region). The first, second and third regions contains approximately 180, 40 and 220 water molecules, respectively.

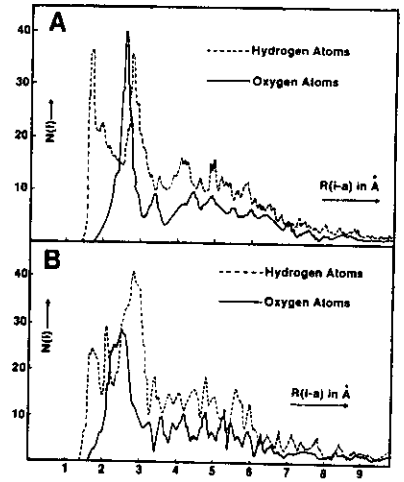


Figure 2. Distribution of oxygen (solid line) and hydrogen atoms (dashed line) internuclear distances from the "a" atom of B-DNA (see text). Insert A for B-DNA, insert B for Na-B-DNA.

The R(H-"a") internuclear distance distribution (see the broken line in Figure 2) is characterized by two very distinct peaks with maxima at 1.8 A and 2.8 A and a distribution of peaks with decreasing intensity for large R(H-"a") values. From these distributions we learn that *most* of the water molecules in the first region are oriented with a hydrogen atom pointing to "a"; since the (H-"a") peak with a maximum at 1.8 A is somewhat broader and a bit more intense than the second one (at 2.8 A), we can also conclude that a small number of water molecules have *both* hydrogen atoms pointing towards B-DNA. The hydrogen and oxygen peaks separa-

tions inform us (2.6 A − 1.8 A = 0.8 A and 2.6 A − 2.8 A = −0.2 A) that, in general, for the first region the hydrogen bond forms structures for the type "a"—H-O.

A detailed analysis of the oxygen atom distribution in the first region, indicates that relatively few water molecules have the oxygen atom (rather than a hydrogen atom) in proximity to an "a" atom. In general, we can associate the water molecules of the first region to the first hydration shell. The water molecules of the second R(O-"a") region are less easily defined.

Let us now analyze the distribution of water molecules in more detail, by considering the probability distribution maps. In Figures 3 and 4 we report the water molecules solvating either the phosphate groups in each of the two helices of B-DNA or the water molecules contained in a disk of about 5 A thickness, (slicing the cylinder perpendicular to the z-axis). In the figures, the water molecules are represented with the new algorithm described above; we shall talk of "water molecule number N", as a short expression to indicate "the ensemble of water molecules that falls within the volume number N, consisting of the previously described three spheres of radius 0.5 A."

In Figure 3 (insert A), we report the ten phosphate groups P1 to P10 of h and the corresponding sugar unit, but not the base-pairs, that are, however, indicated by reporting the terminal nitrogen atom and by using the notation, T1, G2, ..., C10. The outmost circumference has 14.5 A radius; the marks on the figure's frame are at 2 A interval. The water molecules are seen from positive values looking down toward negative z values. Only the water molecules very *near* the phosphate P1 (at z = 18.1 A) to P10 (at z = −12.3 A) are reported. Since the "water molecules" are numbered with an index of increasing value along the z direction (from positive z to negative z), low indices (starting from 1) correspond to water molecules solvating the top of the B-DNA fragment, high indices (approaching 447) correspond to water molecules solvating the bottom of the B-DNA fragment.

Notice in Figure 3 the dual features of the water clusters: not only that the water molecules enclose the PO_4^- group, but also they form *hydrogen bonded filaments* (see for example, the water molecules 322, 230, 332, and 358 in the P8-P9 region). The data in Figure 3 (insert B) provides a view of the water molecules solvating the phosphate groups of the h* strand. In Figure 4 (insert A), we report only one base-pair, the A3-T3* base-pair; the water molecules experience the immediate fields of the G2-C2* and C4-G4* base-pairs (not reported in the Figure in order not to further complicate the drawing). Notice, in addition that in the major groove (Mg) there are ten water molecules (82, 108, 87, 132, 130, 109, 123, 85, 137 and 133) whereas only four water molecules are found in the minor groove (mg) (138, 103, 105, and 99); this situation changes when we extend the major groove and the minor groove volume up to a radius of 14.5 A. The physical reason for these findings is that near to the bases there is more "free" space in Mg than in mg, but further out toward R = 13 or 14 A the field generated by the phosphates in the mg

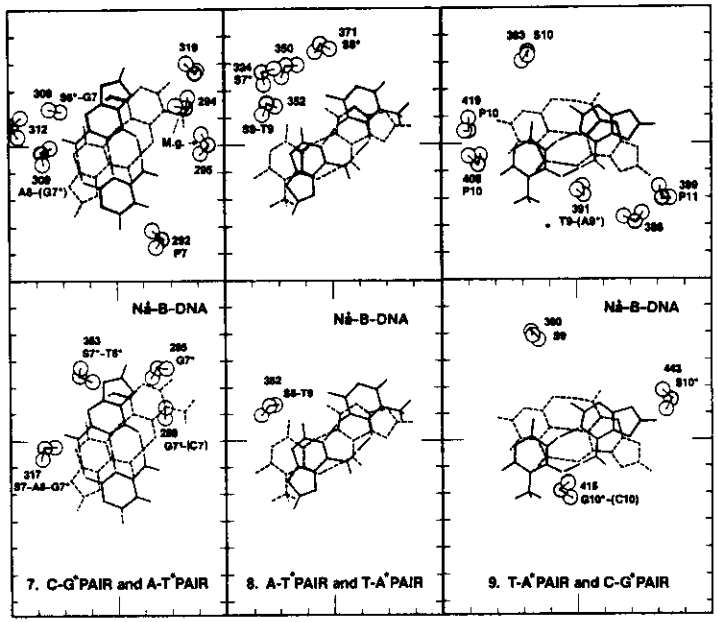


Figure 5. Water molecules in the vicinity of C7-G7*, A8-T8* (insert 7) and A8-T8*, T9-A9* (insert 8) and T9-A9*, C10-G10* (insert 9).

is stronger than the field generated by the phosphates in the Mg. A more quantitative analysis is provided in Tables 1 and 2 of reference 11.

In Figure 5 (top inserts), we report the water molecules solvating three base-pairs. One point should be immediately stressed: we report in these figures not only the water molecules hydrogen bonded to one (or more) bases but also some of those *"nearby"* the bases. For a given water molecule we use the following notation (in addition to the index for the water "volume"): *very strongly* hydrogen bonded water molecules to B-DNA are differentiated from *strongly* hydrogen bonded ones by writing the solvated site either without or within parentheses. Additional figures for the base-pairs are reported in reference 11.

For example, in Figure 5 (insert 7 top) the water molecules reported are in the vicinity of the C(7)-G*(7) and the A(8)-T*(8) pairs; the base-pair (C-G*) is represented with full lines since it is at higher z-value than the lower base-pair (A-T*), represented by dashed lines. The water molecule 292 solvates very strongly the phosphate group in h (P7); water 308 solvates very strongly S6* and also guanine G7 of h. Water 295 is too far from any B-DNA atom to be assigned as solvating a

specific group and therefore, it is labeled only as Mg, namely, one of the molecules in the major groove. Finally water 319 is unlabeled, because it is not strongly hydrogen bonded to any atom of B-DNA. Notice that *often one water binds two bases of two successive base-pairs* (for details see reference 11). In conclusion, we have considered three types of water molecules: *very strongly* hydrogen bonded (to one or more atoms of B-DNA) *strongly* hydrogen bonded, and *weakly* hydrogen bonded. More precisely, water reported as *very strongly* hydrogen bonded are those with an O-"a" internuclear distance not larger than 3.2 A. If the hydrogen bond length is larger, then we classify it as a *weak* hydrogen bond. These criteria are rather restrictive and somewhat arbitrary: the PO_4^- field is very intense and a water molecule "strongly" hydrogen bonded to one of the oxygen atoms in PO_4^-, necessarily strongly feels the field of the remaining atoms in the PO_4^- group. In our classification of "hydrogen bond" we have included, as an additional criterion, the requirement that the overall orientation of a water molecule must be one intuitively reasonable; for example, the oxygen (of H_2O) to oxygen (of PO_4^-) internuclear distance must be larger than the hydrogen (of H_2O) to oxygen (of PO_4^-) distance. In Tables 3 and 4 of reference 11 we have reported in detail the water molecules hydrogen bonded to the sugar units and to the bases, respectively. From these data it is clear that a water molecule bridges very often two bases of two nearby base-pairs, less often two bases of the same base-pair and seldom is *bound* only to one base of a base-pair.

Average Interaction Energies

With these restrictive definitions and considering only *strongly* hydrogen bonded water molecules, we summarize as follows:there are 5.9 water molecules solvating each PO_4^- group, 0.3 water molecules solvating each sugar group, 0.5 water molecules solvating both the sugars and bases (namely, hydrogen bonded bridges between a sugar and a base) and 0.9 water molecules solvating a base. These average values refer only to the first solvation shell, in the strict sense above defined; in this way about 160 water molecules out of a total of 447 are considered. The average water-B-DNA interaction energies (in Kj/mol) are -101.9 ± 5.8, -86.9 ± 4.3, -85.9 ± 4.4 and -63.4 ± 4.2 for the PO_4^-, sugar, sugar and base, base, respectively. The average water-water interaction energies are -6.1 ± 3.8, -12.6 ± 3.7, -12.6 ± 3.5 and -16.6 ± 4.3 Kj/mole for the same groups above listed. We recall[12] that the interaction energy in bulk water is -35.6 Kj/mol.

Trans-groove and Interphosphate Water Filaments

The complexity of the structure of water in the two grooves is evident in Figure 6 (left insert), where we consider the water molecules enclosed in the volume between two co-axial cylinders of radii 8 A and 12 A, respectively. The figure reports the water of only one half of the cylinder volume on a y, z projection. To us the striking feature of this figure is *the clear evidence of hydrogen bonded water filaments from a phosphate group of h to a phosphate groups of h*, spanning the major groove, and of hydrogen bonded water filaments connecting two successive phosphate* groups, from P(i) and P(i+1) in the h (and/or in h*) helix. We have previously commented on these *filaments*[2-5] which were surmised from the iso-energy maps[3] or determined clearly in the A-DNA and B-DNA single helix Monte Carlo simulations;[3,4] however the limited number of water molecules considered in our previous simulations was limiting the validity of our suggestion. We note that this feature—the filament existence—*has been previously encountered in ion-pairs in solution.*[13] *We feel that this feature is basic in any water solution containing ions, and will have profound consequences to the understanding of dynamical and temperature dependent properties of solutions with ions.*

The hydrogen bonds shown in the figure are reported only if the oxygen-oxygen distance (between two waters) is equal or smaller than 3.5 A, and if the oxygen-hydrogen distance is smaller than the corresponding oxygen-oxygen distance. Typical water filaments (see Figure 6) are formed by the waters 310, 323, 343, 377, 388 and 399 linking P7 of h to P11* of h*; or 271, 266, 292, 304, 339, 357, 348, 364 linking P6 of h to P10* of h*; or 220, 257, 288, 299, 328 and 340 linking P5 of h to P9* of h*. These are *transgroove* filaments. Other structured filaments are present and connect a phosphate to a successive phosphate in the same helix. Notice for example, the water molecules for the *inter-phosphate* filaments 402, 384, 364 and 394 from P11* to P10* and 364, 360, 351 from P10* to P9*. In the above examples waters 310 and 399 are *terminal waters of a transgroove* filament; the structure in the *inter-phosphate* filaments is different since the filament 402, 384 and 369 (from P11* to P10*) *continues* with waters 360, 351 leading to P9*. It is stressed that these structures *do not correspond to data obtained by analyzing one or few conformations, but are statistically "stable" and meaningful structures.* It is very tempting to postulate that proton "are transferred preferentially along these filaments." Hence, these filaments are of importance in reactivity studies. *These structures are "dynamical" in the sense that a given structure can evolve into a different structure, at relatively little expense for the total energy of the system.* We refer to a number of early experimental studies[14-25] and to a review paper[26] for comparison to our data.

The experimental sequence of the hydration at different molecular sites[14] is reported to be first at the two free oxygen atoms of PO_4^-, then at the two bonded oxygen atoms of PO_4^-, then at the oxygen atom of the sugar and finally at the bases. These findings are both corroborated by our simulation and complemented with energetic data on the water-site interaction energy. We find that the most attractive site is PO_4^-, followed by the sugar unit and finally by the bases. From our results (and reference 11) one can distinguish between the two free-oxygen atoms and the two bonded oxygen atoms in PO_4^-. However, one must keep in mind the fact that the PO_4^- field is very attractive for water molecules and therefore, a sharp distinction on the attraction for a water molecule from one of the two types of oxygen atoms *independently from the other* is not too meaningful. On the other hand, we recall that the field near the two free oxygen atoms is reinforced by the field of the two bonded oxygen atoms, whereas the field at one of the two bonded oxgyen atoms is weakened by the field of the $-CH_2$ group. These qualitative data are analyzed below in a quantitative way.

The number of water molecules solvating the PO_4^- group at room temperature has been estimated[14] to be between 5 and 6 in DNA neutralized by Na^+ counter-ions. Our simulation yields 5.9 water molecules for B-DNA double helix without counter-ions.

Indirect experimental evidences (angular distribution of near elastic scattering by neutron diffraction[27]) suggests the existence of transgroove water molecules (preferential orientation along the main DNA axis).

Second Approximation: B-DNA, Water Molecules and Na^+ Ions at Fixed Position

In the previous sections we have considered aspects of the solvation problem for a B-DNA double helix fragment. In this section the PO_4^- units have been neutralized with Na^+ counter-ions. As known the double helix either in solution or in fibers, is stabilized by the presence of counter-ions. As previously done, we use the Monte Carlo techniques and *ab-initio derived atom-atom pair potentials*[1,12,13,28-30] and a variable number of water molecules (up to 447) to simulate the hydration at several relative humidity values; all computations are performed at a simulated temperature of 300 K.

For the B-DNA fragment's geometrical characterization, we refer to the previous sections and to the original reference[31]. The Na^+ ion is placed near the free oxygens of PO_4^- at its minimum energy position (the Na^+ atom-atom pair-potentials with model compounds containing the PO_4^- group are obtained in our standard way[29]). *The ion is kept at this fixed position during the Monte Carlo simulation;* this is a non-unreasonable approximation, because the very strong attraction from PO_4^- and Na^+ yields a well pronounced, localized and deep minimum. Since, however, *alternative positions for the Na^+ counter-ions could be considered,* as later discussed in this paper, we present this computation as a *model* study rather than as a *realistic simulation,* which would require the Na^+ counter-ions positions not to be fixed as input but to be determined by the Monte Carlo technique.

In Figure 2 (insert B), we present the histogram for the number of oxygen (or hydrogen) atoms, N(i), (i=O or i=H) having a distance R(i-a) from the nearest atom "a" of Na^+-B-DNA. Comparing this histogram, with the equivalent one for B-DNA (see Fig.2) we notice a broad oxygen atom's peak with its maximum at about R(O-a)=2.5 A and extending up to somewhat more than 3.0 A and three (no longer symmetrically placed) peaks at about R(H-a)=1.8 A, 2.1 A and 2.8 A. Taking as first shell all the water molecules up to about R(i-a)=3.0 A, we conclude that there are more water molecules in Na^+-B-DNA's first shell than in B-DNA. The reason for the hydrogen atom orientation is to be found in the different orientation assumed by a water molecule in the field of Na^+ ions[8,9,13] relative to the orientation in the field of the PO_4^- groups[5]. From Figure 2 we note a richer fine structure in Na^+-B-DNA than in B-DNA, especially in the region between R(i-a)=3.0 A and R(i-a)=9.0 A, suggesting *a more structured water pattern;* in the region R(i-a)=0.0 A to R(i-a)= 3.0 A the integrated area of the distribution N(i) versus R(i-a) is larger than for the

equivalent histogram in B-DNA, giving evidences of more densely packed water. The *ion-induced compression effect* is fully expected on the basis of the energetic differences for a water molecule in Na^+-B-DNA, relative to a water molecule in B-DNA as discussed in Reference 11. Less qualitatively, we obtain for B-DNA

$$n(O) = \int_{R_1}^{R_2} N(O)\, dR = 176 \qquad \text{for R1 = 0. A and R2 = 3.0 A}$$

$$n(H) = \int_{R_1}^{R_2} N(H)\, dR = 373 \qquad \text{for R1 = 0. A and R2 = 3.0 A}$$

and for Na^+-B-DNA

$$n'(O) = \int_{R_1}^{R_2} N(O)\, dR = 211 \qquad \text{for R1 = 0.0 A and R2 = 3.0 A}$$

$$n'(H) = \int_{R_1}^{R_2} N(H)\, dR = 418 \qquad \text{for R1 = 0.0 A and R2 = 3.0 A}$$

Clearly, the corresponding integrals from R1 = 3.0 to R2 = 14.5 A are:

for B-DNA	for Na^+-B-DNA
m(O) = 447 − n(O) = 271	m'(O) = 447 − n'(O) = 236
m(H) = 894 − n(H) = 521	m'(H) = 894 − n'(H) = 476

In conclusion, considering a distance up to R=3.0 A from the nearest atoms of the solute, we find 176 water molecules (and 21 residual hydrogen atoms, belonging to other water molecules) in B-DNA; this number increases to 209 (and two residual oxygen atoms) in Na^+-B-DNA. Alternatively stated, the 22 Na^+ ions in B-DNA have crowded in about 30 additional water molecules into the first solvation shell (or about 1.5 water molecules per Na^+ ion) relative to B-DNA. It is noted that the distance of 3 A, obtained from Fig.2 appears also in the early literature on DNA solvation, particularly in the notable work by Lewin[10].

An obvious consequence (as we shall later discuss in detail) of the ion-induced *compression* effect is that the double helix conformation is *sterically stabilized relative to a less packed situation,* (for example, B-DNA without counter-ions). A second obvious consequence is that the ion-induced compression effect must be some function of the counter-ions interaction energy with water, *thus it will vary from ion to ion.* Experimentally, it is well known that temperature, relative humidity, ionic concentration and specificity are the variables capable to induce conformational transition in DNA (see below).

A complementary global analysis on the structure of water in Na^+-B-DNA and in B-DNA can be obtained by considering the number of water molecules enclosed in the volume between two coaxial cylinders, with the main axis coincident with the B-DNA fragment's main axis (Z direction). We designate as R(i) and R(o) the radius

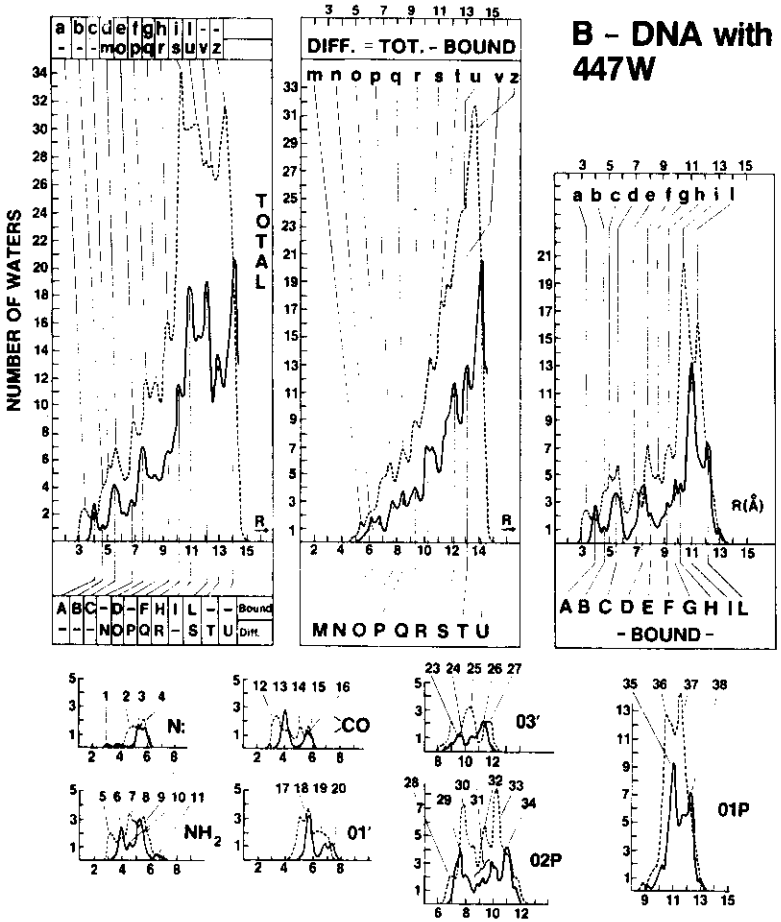
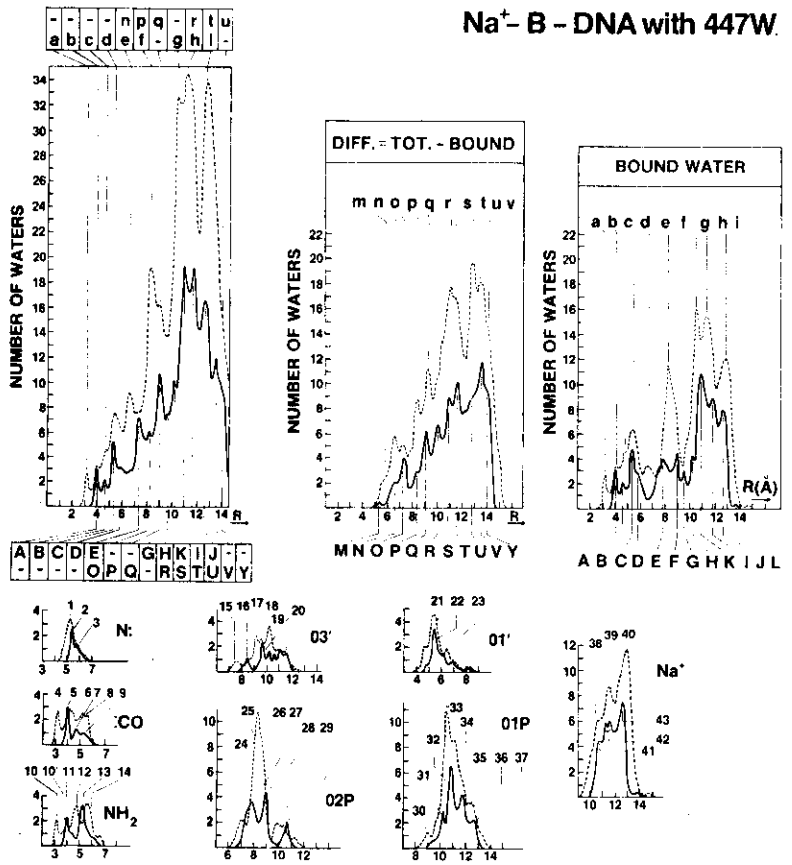


Figure 7. Probabily distributions for water molecule's hydrogen and oxygen atoms as function of R solvating B-DNA and Na^+-B-DNA. The three inserts at the top refer to total, "remainder" and "bound" water distribution (left to right). Figure 7 is continued to right side page.

of the inside and of the outside cylinders, and we select $\Delta R = R(o) - R(i) = 0.2$ A. The resulting diagrams for the hydrogen probability distribution as function of R (dashed line) and for the oxygen atom (full line) are reported in Fig.7 for B-DNA and Na^+-B-DNA. These diagrams can be compared also with the iso-energy contour maps, previously reported for A-DNA, B-DNA double or single helix[2-6,29].



Continuation of Figure 7 from left-hand side page.

In the figures we report the probability distribution for the 447 water molecules (*total*" distribution), a partial distribution related to the water molecules bound to hydrophilic sites ("bound" water molecules distribution) and a second partial distribution *("remainder")* defined as the difference between *"total"* minus *"bound"* distributions.

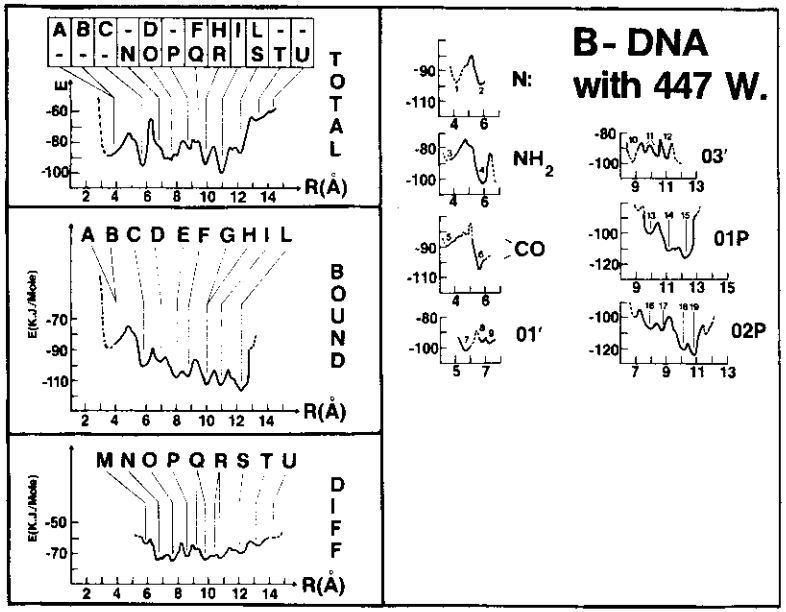
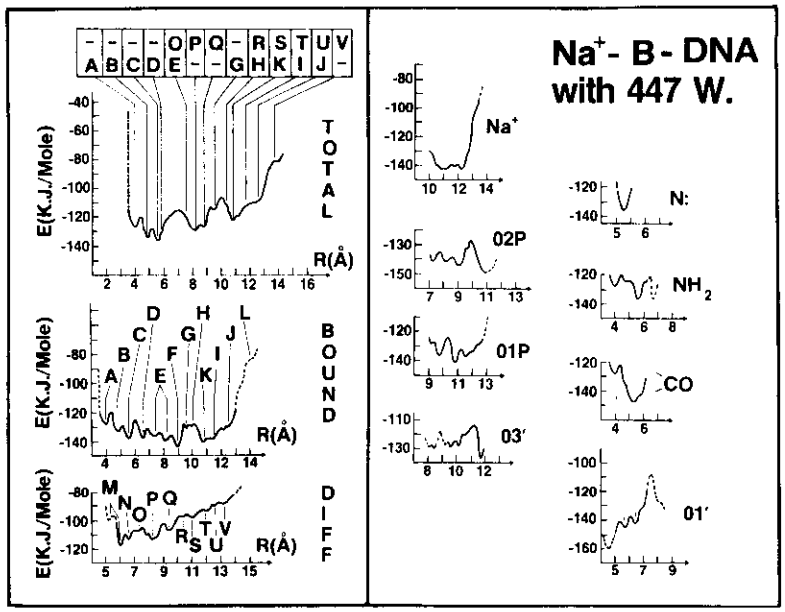


Figure 8. Energy distributions (in Kj/mol) versus R(in A) for the water molecules solvating B-DNA at 300 K. Top left, middle left and bottom left for *"total"* distribution, "bound" water and "remainder"; the other inserts refer to specific "sites" distributions. Figure 8 is continued to right-side page.



Continuation of Figure 8 from left-hand side page.

In Figure 7 we report details of the distribution around specific groups (hereafter, referred as *"sites analyses"*), namely the N:, NH_2, CO sites of the bases, the oxygen O1′ of the sugar, the bound oxygen O3′ and of PO_4^-, the two free oxygens O2P and O1P of PO_4^-, and Na^+, the O5′ has no bound water molecules. In the abscissa we report R in A and in the ordinate we plot the number of oxygen (or hydrogen) atoms contained in the previously described volume element defined by R(i) and R(i)+0.2 A. Since a given water molecule (as previously pointed out) can belong to the first solvation shell of more than one atom (of the solute) the distributions of the water molecules for the N:, NH_2, CO, O1′, O3′, O2P, O1P and Na^+ groups are *not* additive.

In Figure 7 the oxygen atom peaks are identified with capital letters, A, B, ... U and the hydrogen atoms peaks are identified with lower case letters a, b, ... z. For the site's diagrams the peaks are indicated with numerals.

Before analyzing the distribution data in Figure 7, we explain the main features of Figure 8. For each probability distribution (Figure 7) there is a corresponding energy distribution, reported in Figure 8. The energy units are Kj/mol. The minima in the energy distribution are given by capital letters for the *"total"*, *"bound"* and *"remainder"* waters, by numerical indices in the "sites" analyses. The energy distribution are given as full lines or as dashed lines depending on the statistical significance of the energy. For example, in Figure 8, the energy distribution for the N: atoms (bases) shows two minima, designated as 1 and 2; the corresponding probability density distribution (Fig.7) is low in the interval R=2 A to R=4 A, therefore the corresponding energy distribution is statistically not too meaningful, and is reported as a dotted line rather than a full line.

The analyses of Figures 7 and 8 is facilitated by recalling the R value of few important groups or atoms of the DNA fragments (where R is the distance of a given DNA atom from the long axis of DNA). For *adenine* the two hydrogen atoms of NH_2, are at R=1.8 A and R=2.8 A and the lone pair nitrogen atoms are at R=3.2 A (N3) and R=4.0 A(N7). For *cytosine* the two hydrogen atoms of NH_2 are at R=2.1 A and R=3.7 A, and the oxygen atom at R=3.7 A(O2). For *guanine* the two hydrogen atoms of NH_2 are at R=3.0 A and R=4.0 A, the two lone pair nitrogen are at R=3.3 A (N3) and R=3.9 A(N7) and the oxygen atom at R=1.7 A(06). For *thymine* one oxygen atom is at R=2.8 A(04) and the other is at R=3.6 A(02). These

are not the possible bonding sites for the single bases or the base-pairs in solution, but for the base-pair in B-DNA double helix. As known, more sites are known in solution[2] for separated base-pairs, but these are not present in B-DNA mainly due to steric hindrance because of the base-pair stacking and the presence of the sugar groups. For the sugar the oxygen atom O1′ is at R=6.2 A. For the *phosphate groups* the two free oxygen atoms are at R=10.2 A (O1P) and R=9.1 A(O2P), whereas the two bound oxygen atoms are R=6.7 A(O5′) and at R=8.8 A(O3′); we recall that O5′ is near the CH_2 hydrophobic group. The *sodium* ions are placed as previously discussed at R=10.7 A.

With all the above in mind, we can now analyze the data of Figures 7 and 8. Let us start by comparing the *"total"* distributions of the 447 water molecules in B-DNA and in Na^+-B-DNA (top left inserts in Figure 7). We notice a set of peaks for the hydrogen and for the oxygen atoms; the two higher peaks for the oxygens (peaks K and I) at R=10.8 A and R=11.7 A for Na^+-B-DNA can be compared to the highest peak, U at R=14.2 A, in B-DNA. The shift confirms the ion induced *compression effect*, previously mentioned in analyzing Figure 2 (insert B). Equally evident is the main overall feature in the hydrogen distribution (relative to the oxygen distribution); the hydrogens are shifted toward smaller R values in B-DNA than in Na^+-B-DNA; in addition the B-DNA and the Na^+-B-DNA patterns of peaks are different. A peak by peak identification can not be made without reference to the distributions of the *"remainder"* water (Figure 7, top middle inserts) and of the *"bound"* water molecules (Figure 7, top right inserts). By integration of the "bound" distribution we find that in B-DNA 157±2 water molecules out of the 447 are "bound" water. The "remainder" water molecules population increases with R (as fully expected), with a characteristic pattern (peaks M to U in B-DNA, M to Y in Na^+-B-DNA). By comparing Na^+-B-DNA with B-DNA we notice that the "remainder" remarkably shows the ion induced compression effect. This is to be expected; indeed these molecules can be compressed more readily than the bound molecules, the latter being trapped at the sites, that is, at the most intensive and attractive field region, as shown by the energy distributions for the water molecules in the "remainder" relative to those in the "bound" (see also Figs.11 and 12, the three inserts at the left side).

The presence of the counter-ion brings about another global effect: in Na^+-B-DNA the *bound* water molecules have approximately a constant energy attraction from R=4.0 A to R=12.0 A, whereas in B-DNA the bound water molecules are more attracted at R=13.0 than at lower values of R. *Thus, the counter-ion brings about an over-all increase in the attraction for all the water molecules, and this effect is more prominent for the water molecules near the base-pairs than for those at large R values.* Another important feature is that the "remainder" water molecules show the lowest energy at R=6.0 in Na^+-B-DNA, (minima M and N) whereas the lowest energy is between R=6.0 A and R=11.0 A in B-DNA (minima N, P, Q, and R). This observation brings about at least two conclusions. *First, a molecule* (for example, a carcinogenic or an anti-carcinogenic molecule with polar groups) *is expected to be "pressed" toward the base-pair by the field of the counter-ions, displacing*

groove's water molecules. Second, theoretical computations on carcinogenic or anti-carcinogenic compounds performed on models where only a single base or a base-pair are considered, rather than a full fragment of DNA with its counter-ions might be not too relevant to problems related to the DNA interaction with such compounds, unless, case by case, it has been quantitatively demonstrated that the solvent and counter-ion effects are small relative to the computed interactions between DNA and the chosen compound (at the temperature in consideration).

Let us continue with the analyses of the data in Figures 7 and 8; in this section we shall focus on the *"bound"* water molecules. To understand the density distribution peak by peak and the energy minima, we have to consider the *"sites* analyses" reported in Figure 7 and 8. Let us start with B-DNA (Figure 7, top right insert and the seven small insert at the bottom). *Peak A* is formed by peak 6 of NH_2 and by peak 13 of CO; *peak B* is formed by peak 11 of NH_2; *peak C* is formed by peak 3 of N:, peak 8 of NH_2, peak 18 of O1′ and peak 16 of CO; *peak D* is formed by peak 29 of O2P, and peak 20 of O1′; *peak E* is formed by the right shoulder of peak 29 of O2P (not identified by a number); *peak F* is formed by peak 24 of O3′, by peak 31 of O2P (the peaks at R=8.8 A, 9.3 A, 9.8 A and 10.2 A have been collectively labelled as peak 31 in O2P) and by the beginning of peak 35 of O1P; *peak G* is formed mainly

Table I

Solvation of B-DNA without and with counter-ions at T=300 K: water's population bound to sites and its interaction energy (Kj/mol)*

#	Site's Name	B-DNA		Na^+-B-DNA		# of Sites
		Waters	Energy	Waters	Energy	
1	O1P	62.29	−108.18	47.47	−133.54	19
2	O2P	46.44	−111.12	35.91	−141.97	19
3	O5′	1.59	−90.40	0.95	−131.25	19
4	O3′	16.49	−91.09	17.56	−124.92	19
5	Na^+	-----	-----	65.87	−134.73	19
6	O1′	15.29	−98.68	16.02	−139.87	19
7	NH_2 (A)	5.53	−83.93	4.50	−126.93	5
8	NH_2 (C)	9.84	−79.96	10.56	−120.87	5
9	NH_2 (G)	4.86	−93.80	5.58	−141.66	5
10	N3 (A)	2.27	−85.10	1.78	−136.07	5
11	N7 (A)	0.95	−84.63	1.34	−121.55	5
12	N3 (G)	2.54	−98.23	2.68	−139.84	5
13	N7 (G)	1.16	−83.46	2.45	−120.66	5
14	O2 (T)	1.98	−108.00	2.96	−151.03	5
15	O4 (T)	4.58	−80.10	4.87	−125.50	5
16	O2 (C)	2.00	−98.51	2.00	−142.77	5
17	O6 (G)	5.16	−86.58	5.26	−126.29	5
18	Boundary (type 1)	14.99	−106.94	11.62	−130.72	---
19	Boundary (type 2)	0.00	-----	7.78	−130.84	---
20	Total	158.87	−103.87	192.52	−133.75	

*In the last column we report the number of times a given site is present in the fragment (excluding the top and bottom which are considered in the boundaries).

Table II
Solvation of B-DNA without and with counter-ions at T=300 K:
population of water bound to groups of sites and its interaction energy (Kj/mol)*.

#	Groups of Sites	B-DNA		Na+-B-DNA	
		Waters	Energy	Waters	Energy
1	O1P, O2P, 05', 03'	113.36	−107.35	93.54	−135.84
2	O1P, 02P, 05', 03', Na+	-----	-----	141.70	−134.93
3	NH_2 (A, C, G)	17.56	−84.39	18.72	−127.56
4	N: (A, G)	6.92	−89.57	8.25	−130.36
5	O (C, G, T)	11.32	−90.99	12.96	−134.09
6	Adenine	8.45	−83.90	7.62	−128.12
7	Cytosine	11.83	−83.09	12.56	−124.36
8	Guanine	13.25	−91.00	14.81	−132.78
9	Thymine	6.56	−88.52	7.83	−135.15
10	A-T	12.56	−86.60	13.87	−132.55
11	G-C	21.14	−87.15	22.76	−129.18
12	A-T, G-C	24.56	−87.04	30.04	−130.16
13	A-T, G-C, O1'	31.15	−89.20	35.56	−130.17
14	(ph-S) (base-pair) (S-ph)	15.05 ± 0.10		19.00 ± 0.10	

*See Table I (last column) for number of sites (and, therefore, of groups of sites).

by peak 31 of O2P, peak 24 of O3' and peak 35 of O1P; *peak H* is formed by peak 35 of O1P, by peak 31 of O2P, and by peak 26 of O3'; *peak I* is formed by peak 35 of O1P (the maximum of this peak), by peak 34 of O2P and by peak 26 of O3'; finally, *peak L* is formed by peak 38 of O1P (the low sub peak of L corresponds to the low sub-peak of peak 38 of O1P (at R=13.0 A)). Thus each feature of the "bound" water distribution is now *fully identified.*

In Tables I to IV we condense the finding discussed up to now. In these tables additional "site" decompositions are presented relative to those given in Figure 7, mainly to further characterize those sites known to be of basic importance in problems related to carcinogenic activity and intercalations. We have *excluded* from the analyses the top and the bottom base-pairs (A0-T0* and G11-G11*) and the connecting three-phosphate groups and five sugar units. The exclusion, was made in order to delete possibly spurious data due to boundary condition artifacts. The water molecules at the top and bottom's boundaries are designated as *"Boundary" type 1 or type 2* depending if the binding is of type "a"-H-O or "a"-O. Table I considers water molecules "bound" to specific sites (like O1P or O2P); Table II considers groups of sites (like all the oxygen atoms in PO_4^-); Table III considers not only "bound" water molecules, but all those forming the first hydration shell; Table IV considers the water molecules around hydrophobic groups containing H atoms (-CH, CH_2 and CH_3) and is given to allow further interpretation on infrared or Raman and nmr studies at the carbon atoms. *The information provided by these self-explaining tables compared to the experimental data accumulated in the last twenty years can be taken as an indication of the evolution in simulation techniques.*

Table III
Solvation of B-DNA without and with counter-ions at T=300 K:
population of water and its interaction energy (Kj/mol)
in the first Solvation shell.

#	"Regions" within first Solvation Shell*	B-DNA		Na-B-DNA	
		Waters	Energy	Waters	Energy
1	O1P, O2P, O5%, O3', CH_2	136.78	−104.51	146.10	−133.39
2	O1P, O2P, O5', O3', CH_2, Na+	-----	-----	170.09	−133.38
3	O1'	15.36	−98.04	16.21	−139.90
4	sugar (C_4OH_5)	83.83	−97.79	100.53	−134.21
5	NH_2 (A, C, G)	19.97	−85.93	20.77	−127.85
6	N: (A, G)	7.08	−89.71	8.66	−129.92
7	O (C, G, T)	11.32	−90.99	12.96	−134.09
8	Adenine	11.81	−86.63	13.93	−133.89
9	Cytosine	19.56	−87.33	18.11	−124.48
10	Guanine	19.54	−95.24	20.47	−135.05
11	Thymine	15.63	−87.14	15.33	−132.18
12	A-T	24.37	−87.30	23.93	−132.47
13	G-C	34.84	−91.73	33.40	−130.46
14	A-T, G-C	47.80	−91.14	50.18	−131.08
15	A-T, G-C, (C_4OH_5)	106.74	−94.32	122.72	−121.94
16	Total for first solvation shell	185.31	−99.74	223.39	−130.63
17	Total for Grooves	261.70	−64.39	223.62	−91.74
18	Total (water-water & water B-DNA)	447.00	−79.36	447.00	−111.30
—	(Water-water + total interaction)	−15.61 ± 0.05		−22.0 ± 0.05	
—	(Water-B-DNA + total interaction)	−63.75 ± 0.10		−89.3 ± 0.1	

*See the last column of Table I.

Table IV
Water molecules perturbing hydrogen atoms (hydrophobic) in B-DNA
without and with counter-ions at T=300 K*.

Group	Atoms of group	B-DNA		Na+-B-DNA	
		Waters	Energy*	Waters	Energy*
CH_2	H5'1, H5'2	43.32	−95.12	46.95	−128.30
C_4OH_5	H3, H4	55.97	−95.80	64.17	−131.06
"	H2'1, H2'2	22.13	−99.89	29.69	−137.11
"	H1'	12.44	−98.98	10.80	−139.79
Bases	H in A**	3.24	−92.57	4.63	−144.97
"	H in C**	9.56	−89.08	8.23	−123.30
"	H in G**	5.00	−104.24	7.09	−134.98
"	H in T**	10.04	−95.12	9.38	−128.49

*Kj/mol.
**The hydrogen atoms either involved in the base-base hydrogen-bonds or in the NH_2 groups are *not* considered in this Table.

The "sharing" of water molecules at different sites has been pointed out previously in the single-helix (A-DNA and B-DNA) solvation studies[3,4] as well by Lewin[10]. In this paper a more definite answer is provided. For example the sites at the bases (lines 6 to 9 in Table II) are cumulatively solvated by 24.56 water molecules, but considering the sites, one by one, 40.09 molecules of water are involved, namely about 40% (15.53) of the water molecules are "shared". The "sharing" percent in the water bound to the phosphate sites (lines 7 to 10 and composite analyses at line 11) is rather small and we interpret this fact as a consequence of our stringent criteria for the thresholds T(a-H) and T(a-O); relaxing these values we would include "second" solvation shells, where the "sharing" is even more prominent (see Ref.4 and 9).

Structure of Water at High Relative Humidity: "Average Configuration" Representation

Let us now analyze the water structure making use of the *"average configuration"* data. In Figures 4 (insert B) and 9, we analyze the water molecules either included in a disk volume 4.4 A thick (from Z=12.1 A to Z=7.7 A, see also Figure 4 (insert A) or the water molecules hydrogen bonded to the h* helix (see also Figure 3 (insert B).

In Figure 4 (insert B), we report the third base-pair A3-T3* and those water molecules *fully inside the disk*. Several water molecules below discussed are only partly within the disk and *are not reported in the figure.* To simplify the figure we have omitted the G2-C2* base-pair immediately above A3-T3* and the C4-G4* base-pair, immediately below. There are four phosphate groups in this disk, P3, P4 (top of figure) and P3*, P4* (bottom of figure). The four Na^+ ions are indicated with a full dot connected to the free oxygen atoms of PO_4^- by dashed lines. The P3 phosphate group is hydrogen bonded to the water molecules 60, 100, 124, 132 and 133; the P4 phosphate group is hydrogen bonded to the water molecules 95, 121, 160, 172 and 185; the P3* phosphate group is hydrogen bonded to the water molecules 73, 85, 105, 128 and 147; the P4* phosphate group is hydrogen bonded to the water molecules 109, 112, 131 and 156. The Na^+ ions are solvated as follows:$Na^+(3)$ is hydrogen bonded to the water molecules 83, 97 and 126; $Na^+(4)$ is hydrogen bonded to the water molecules 122, 159 and 160; $Na^+(3)$* is hydrogen bonded to the water molecules 74, 93 and 128; $Na^+(4)$* is hydrogen bonded to the water molecules 123, 131, 167 and 169. The water molecules 128, 131 and 160 establish a hydrogen bond to both the phosphate and the Na^+ ion. Concerning the base-pairs, the water molecule 91 is hydrogen bonded to G2, T3* and C2*; 111 is hydrogen bonded to A3; 140 is hydrogen bonded to T3* and G4*. The remaining water molecules of Figure 13 are either second or third shell solvation and therefore fill the major or the minor groove.

In Figure 9, we consider the water molecules hydrogen bonded to the Na^+-PO_4^- groups in the h* helix; in the figure we have indicated the sugar residues and the N1 or N9 atoms of the bases. Comparing this figure with Figure 3 the denser water packing in Na^+-B-DNA relative to B-DNA is apparent.

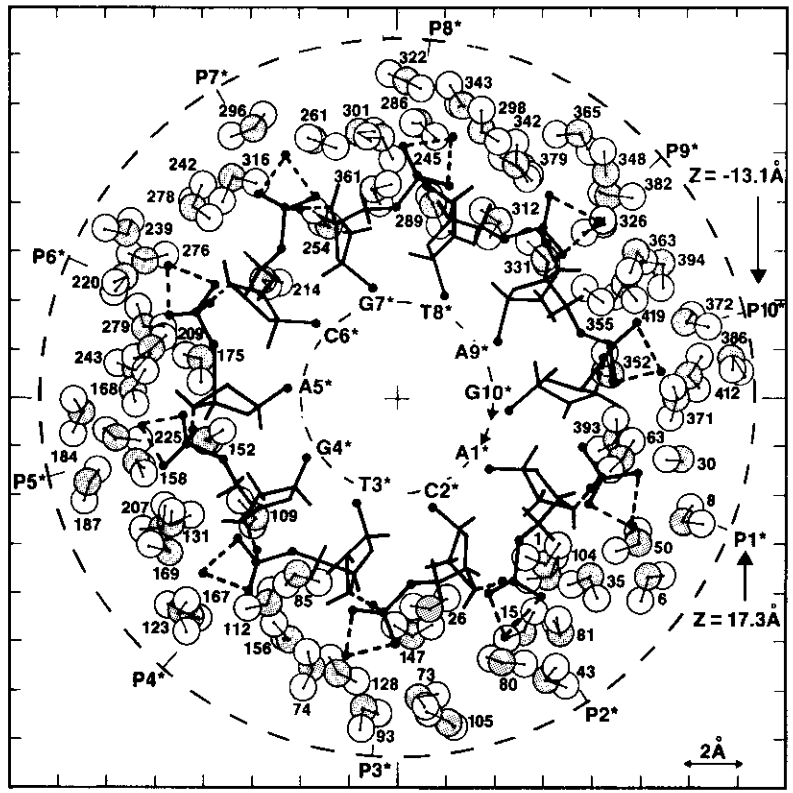


Figure 9. Water molecules solvating the sodium-phosphate groups of the h* strand of Na^+-B-DNA; "average" configuration.

In Figure 5 (bottom inserts) we report the water molecules "bound" to three base-pairs or in its immediate neighborhood. Additional figures for the base-pair sequence of the B-DNA fragment are reported in reference 32. One base-pair (the one at higher Z) is presented more markedly than the second, which lies immediately below (looking from the Z axis). From this figure, the average position and orientation of the water molecule solvating the base-pair can be obtained. The structure accepted by the water molecules represents the effect of the totality of the forces acting on each water molecule, namely the effect of the interaction energy between a given water molecule and the rest of the solute-solvent system. By shifting the stress from *forces* to interaction energy we implicitly restrict ourself to a static representation (Monte Carlo) and we neglect the dynamic representation (molecular dynamics).

Today's literature on "theoretical" studies for solvation of biomolecules often neglects temperature averaging and makes use of a "single configuration" despite the availability of Monte Carlo techniques, proposed about twenty years ago. In addition, in today's literature often too small a fragment and too few water molecules are considered. For example, B. Pullman and his group[33] have analyzed the solvation of a three base-pair fragment of B-DNA. The water molecules were placed one after the other at one half of the fragment, *not all at once.* The implicit assumption of these authors is that the position for one water molecule does not depend on the positions of the *following* water molecules, but mainly depends on the positions of the *previously* placed water molecules. As known (see Ref. 34) this assumption is incorrect and in the Monte Carlo technique much effort goes into erasing the memory of the starting configuration. Since from experimental data it was known that 4 to 6 water molecules solvate a phosphate group, 17 water molecules were selected[33]; five were placed at each phosphate group and the remaining two were placed at the three base-pairs. This starting configuration unfortunately biases the final geometry output. Therefore, the simulation[33] reports mainly "boundary effects" (since the fragment is too small), at low-intermediate relative humidity (since 17 water molecules are too few to describe high humidity) and with temperature neglection. This type of computational study, if improved, yields structural data not dissimilar from those obtained using "sticks and balls" models, as clearly demonstrated in the early work by Lewin.[10]

Structure of Water in the Grooves

We recall that in the first approximation we have commented on the *intra-phosphate water filaments* (connecting P(n) to P(n+1) or P*(n) to P*(n+1)) and *inter-phosphate or "trans-groove" water filaments* (connecting P(n) to P*(m)). In Na^+-B-DNA the ion-induced compression effect brings about an added feature relative to B-DNA, namely we observe hydrogen-bond cross-linking between inter-phosphate water filaments.

In Figure 6 (right insert), we examine the structure of water in the minor and in the major grooves selecting waters with $8 \leq R \leq 13$ A. The water molecules are projected onto the Y-Z plane, namely a plane containing the long axis (Z axis) of DNA (see Fig. 10); we have selected those water molecules with positive value for the X coordinate of the oxygen atom. The figure is seen from the +X direction; the five atoms forming a phosphate group are reported; full dots represent the two bound oxygen atoms O3′ and O5′, whereas open dots represent the two free oxygen atoms O1P and O2P. The Na^+ ion is represented by a full dot. The space enclosed by two lines passing through P3, P4, P5, P6 and P7 and through P8*, P9*, P10*, P11* corresponds to the major groove; above and below it are portions of the minor groove. Dotted lines between atoms of water molecules indicates a water-water hydrogen bond.

For Na^+-B-DNA the solvation pattern can be described either in term of *cross-linked* inter-phosphate water filaments or in terms of local *cyclic structures* com-

posed of about 5±2 water molecules. For example, the water molecules 293, 321, 330, 359, 371 can be considered as a inter-phosphate filament; a second filament is formed by the water molecules 293, 327, 338, 350, 354, 375, 403. The cross-linking water molecules for these two inter-phosphate filaments are water 354 and 375. Alternatively, cyclic structure representations are given, for example, by the water molecules 327, 321, 330, 359, 375, 354, 350, 338 or 375, 359, 371, 393, 403. At this stage we do not see any reason to prefer the cross-linked filament "model" over the cyclic structure "model".

The above "low" resolution structure of water molecules in the grooves is now complemented with the "high" resolution obtained not from the "average" configuration, but from a full statistical analysis on the 2×10^6 Monte Carlo configurations.

First we partition the water molecules into those belonging either to the major or to the minor groove. The cylinder containing the solvent-solute system is subdivided into sectors (see Fig. 10, left side insert). For a full B-DNA turn there are N sectors of angle α and height (along Z) equal to 13.2 A for the minor groove and 20.08 A for the major groove; for a full B-DNA turn there are $360/\alpha = N$ sectors (in the figure we have represented four sectors in the minor groove). We select a clockwise rotation for α (with $\alpha = 0$, $y = 0$ and x positive, see left side of Fig. 10). Finally, to increase the resolution in locating a water molecule in the grooves, each sector is partitioned into four sub-volumes, as indicated at the bottom left insert of Fig. 10, by subdividing R into four segments, namely $0.0A \leq R < 4.0$. A, $4.0A \leq R < 8.0A$, $8.0A \leq R < 12.0A$ and $12.0A \leq R < R(max) = 14.5$ A. In the right side insert of Fig. 10 we report the phosphorous atoms of the h and h* strands; these atoms are on a cylindrical surface and have been projected onto XZ plane. When α increases from 0° to 180° (or 180° to 360°), we consider all those water molecules with negative (or positive) value for Y. We count the water molecules which fell into one of the four sub-volumes of a given sector selecting those left out from the previous analyses performed for the first solvation shell (therefore, only true "groove" molecules are

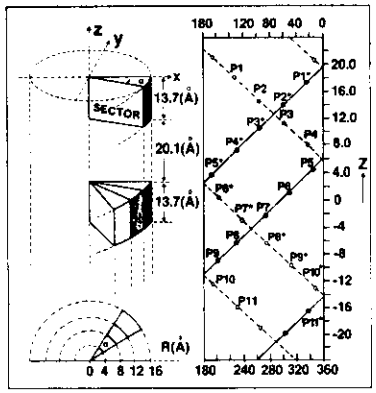


Figure 10. Grooves analyses: sectors and sub-volumes (left), major and minor (dashed) grooves (right); the bottom scale is seen from the -Y axis, the upper scale from the Y axis.

considered). After several trials we have selected $\alpha = 4°$ as the value yielding a clear graphical representation which is reported for the major groove in Fig. 11 (top insert) (the values reported on the abscissa follow the definition used in Fig. 10); the dashed vertical lines identify the position for the P* atoms (of the h* strand) and the marks at the top of the figure identify the position for the P atoms (of the h strand). In the figure we compare Na^+-B-DNA with B-DNA. In the region 0. $A \leq R < 4.0A$ there is no groove water, as previously pointed out for example, by considering iso-energy maps[2,6,29].

The number of water molecules (Nw) and the average interaction energy per water molecule, E (in Kj/mol) varies with R as reported below:

Limits	Nw	E	Groove	Solute
$4 \leq R < 8$	0.38	−125.38	minor	Na^+-B-DNA
$4 \leq R < 8$	15.82	−104.25	major	Na^+-B-DNA
$4 \leq R < 8$	0.0	0.0	minor	B-DNA
$4 \leq R < 8$	9.22	−66.93	major	B-DNA
$8 \leq R < 12$	19.70	−98.52	minor	Na^+-B-DNA
$8 \leq R < 12$	74.02	−97.22	major	Na^+-B-DNA
$8 \leq R < 12$	18.94	−72.47	minor	B-DNA
$8 \leq R < 12$	64.87	−66.77	major	B-DNA
$12 \leq R \leq R(MAX)$	49.77	−88.54	minor	Na^+-B-DNA
$12 \leq R \leq R(MAX)$	64.41	−82.58	major	Na^+-B-DNA
$12 \leq R \leq R(MAX)$	64.92	−63.75	minor	B-DNA
$12 \leq R \leq R(MAX)$	102.95	−61.76	major	B-DNA

From the above data we learn new feature on the packing of the water molecules in the grooves.

The interaction energy in the minor groove is larger than in the major groove (as expected because the attractive sites are nearer in the minor groove). For B-DNA the interaction energy is nearly constant in the major groove, as previously predicted[2,5] but it increases sharply by decreasing R for Na^+-B-DNA. *Thus the grooves can be seen as channels to convey reactants to the base-pairs;* the "flow" in such channels depends both on the counter-ion position and specificity, and on sterical factors of the reactant.

In Fig. 11 a more detailed representation of the "groove" water molecules in the major groove is given. An equivalent diagram for the minor groove shows the same periodicity of peaks but with lower intensity; these peaks represent the "intra" phosphate filaments previously discussed. The gross characterization of the intensity distribution given in Fig. 11 is the nearly periodic existence of peaks, approximately evenly spaced ($\Delta\alpha = 10°$), well developed in the regions $8 \leq R < 12$ and $21 \leq R \leq R(max)$ and weakly as well as infrequently developed in the region $4A \leq R \leq 8$. The peak pattern *clearly points out the existence of a structural organization which even if temperature dependent, is not destroyed by thermal motion at*

room temperature. The intensity of the peaks (i. e. , the number of water molecules bound to the phosphates) adds up to "filaments" of about 5 to 7 water molecules.

The number of water molecules, Nw, and the average interaction energy per water molecule, E, with consideration given not only to the groove's water but also to the

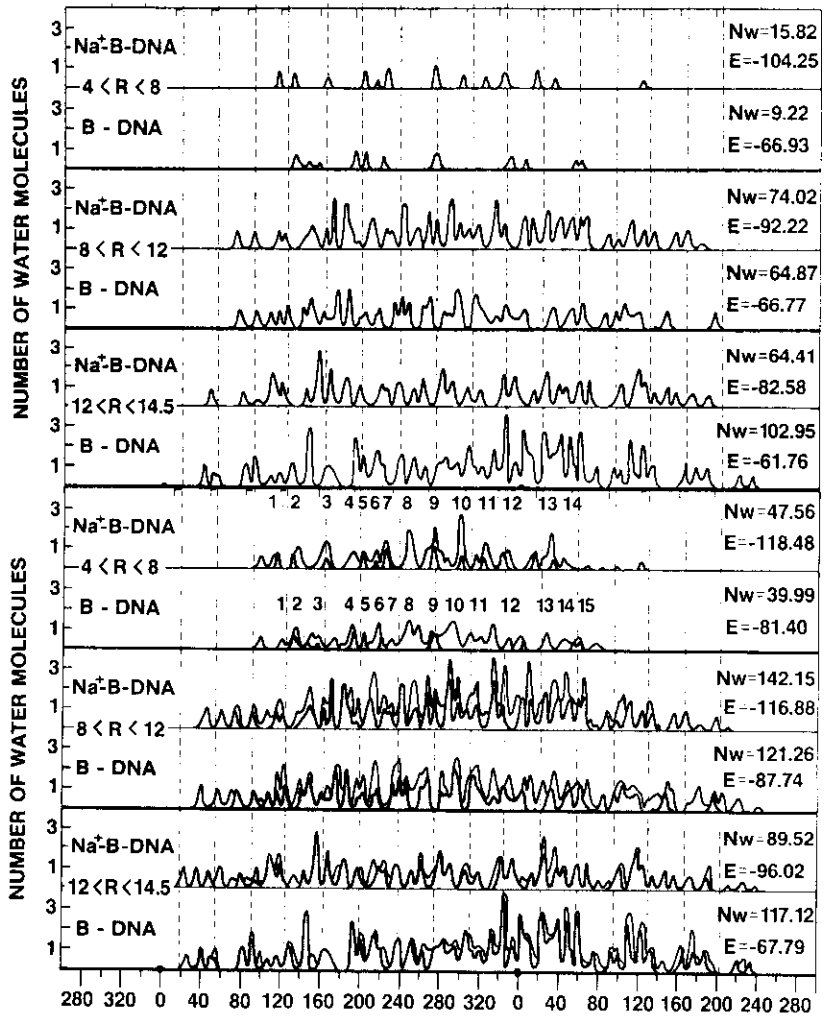


Figure 11. Probability distributions for water molecules in B-DNA and Na^+-B-DNA major groove (top) first solvation shell and groove water molecules (bottom).

first solvation shell water molecules, are partitioned into the sector's sub-volumes are as reported below:

Limits	Nw	E	Groove	Solute
0≤R<4	0.0	0.0	minor	Na^+-B-DNA
0≤R<4	1.15	−125.25	major	Na^+-B-DNA
0≤R<4	0.0	0.0	minor	B-DNA
0≤R<4	1.92	−87.10	major	B-DNA
4≤R<8	20.25	−137.12	minor	Na^+-B-DNA
4≤R<8	47.56	−118.48	major	Na^+-B-DNA
4≤R<8	19.79	−94.56	minor	B-DNA
4≤R<8	39.99	−81.40	major	B-DNA
8≤R<12	80.83	−117.46	minor	Na^+-B-DNA
8≤R<12	142.15	−116.88	major	Na^+-B-DNA
8≤R<12	74.36	−87.54	minor	B-DNA
8≤R<12	121.26	−87.74	major	B-DNA
12≤R≤R (MAX)	65.56	−98.90	minor	Na^+-B-DNA
12≤R≤R (MAX)	89.52	−96.02	major	Na^+-B-DNA
12≤R≤R (MAX)	72.56	−68.33	minor	B-DNA
12≤R≤R (MAX)	117.12	−67.96	major	B-DNA

One can notice that the distinction between the major and minor groove, extends outside the dimension of the solute. Remembering that the Na^+ ions are at the periphery of B-DNA (at R=10 A), the groove structure extends into the liquid surrounding DNA. We shall call "classical groove volume" the one up to R=12. A, and *"extended groove volume"* the one starting at R=12. A. The latter ends when E=−36 Kj/mol[12], *namely when the value of R for the water is equal to about 20 to 25 A.* In Fig. 11 (bottom insert) we report the water molecules present in the major groove, considering not only "groove" water but also first solvation shell water molecules. The peaks presented as dashed areas correspond to *groove* water molecules; those presented as a line-contour correspond to *groove* and *first solvation* shell water molecules. To facilitate the understanding of the figure, some of the peaks are labelled with numerals. In the first sub-volumes (O.<R<4.A), most of the water molecules are present as first solvation shell rather than as groove waters. Peaks 1, 2, 3, 5, 6, 7, 9, 10, 11 and 12 have a groove's water contribution, which is sometimes very small (for example, peak number 10); no *groove* contribution is sometime present (for example, peaks 4, 8 and 13). In Na^+-B-DNA the groove contribution is larger than in B-DNA, because of the ion compression effect. The NH_2 groups for A, G and C are in the vicinity of the peaks 1, 2, 3, 8, 9, 13, 14 and 15; the O4(T) and the O6(G) atoms are in the vicinity of the peaks 4, 5, 7, 11 and 12; nitrogen lone pairs (N7 of G and A) are in the vicinity of the peaks 4, 6, 8, 9, 10, 12 and 13. Rotations of water molecules (and thermal effects) are expected to be appreciable mainly for the water molecules corresponding to the dashed area.

In the sub-volumes from R=4 A to R=8 A each first solvation peak is enhanced by groove water contributions (exceptions are at the boundaries). The peaks intensity

is higher for Na^+-B-DNA than for B-DNA, as expected. *The emerging picture is that a water filament starts as bound water, continues as groove water to end as bound water, in agreement with the previous findings.*

In the sub-volumes from R=12 A to R=14 A the groove water is responsible for most of the intensity of the peaks, especially in Na^+-B-DNA.

Water Solvation of DNA at Low and Medium Relative Humidity

As known (see Refs. 14, 15, 16 and 26) in adsorption-desorption experiments, the isotherm presents a characteristic growth curve. The experiments are generally performed on DNA fibers, and in these conditions there are several parameters to keep in mind, even at constant temperature, constant relative humidity and constant ionic concentration. By decreasing (or increasing) the number of water molecules, there is a reorganization in the entire water system. In addition the relative position of the DNA units in the fiber varies (a swallowing of the fiber is observed). Moreover, concomitant with the reorganization of the water and of the DNA units within the fiber, conformational variation in the DNA can occur. *Finally*, fiber diffraction data are limited to about 3 A resolution. An advantage in simulating the water structure around a fixed-geometry Na^+-B-DNA fragment is that the energetic and structural variations of the water system can be followed, without the worry of the notable additional complications associated with the interactions of one DNA molecule (and its solvent) with other DNA molecules in the fiber.

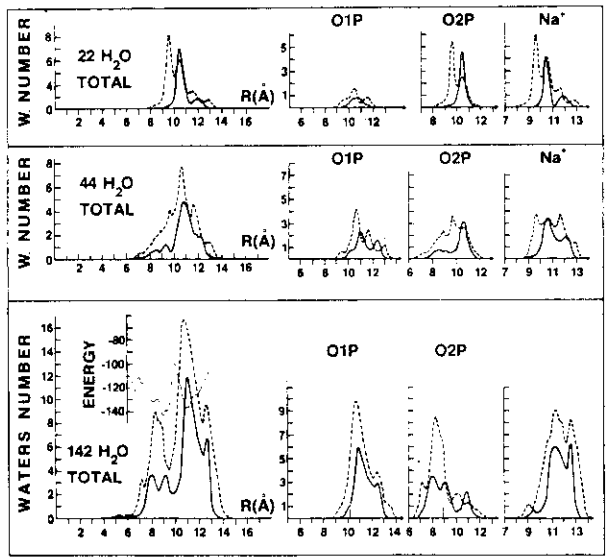


Figure 12. Probability distributions for 22, 44 and 142 water molecules in Na^+-B-DNA.

We precede the following analyses by reporting average distributions computed by adding to the Na^+-B-DNA fragment either 22, 44, 142 or 257 water molecules. The results are reported in full detail in Tables XI to XIV of reference 32 and are summarized here in Figures 12 and 13.

For 22 waters the sites interested are Na^+, O1P and O2P; one water molecule is bound to each PO′-Na^+ unit. A water molecule is bound with about equal probability to Na^+ and O2P, but less to O1P. The O5′ and O3′ site "do not participate" (the analyses on O1P and O2P yields a sum of 17.80 water molecules); by considering also O5′ and O3′ and the CH_2 group, this value does not change. Essentially two distributions are found, one with a water bound both to Na^+ and O2P, and a second, less probable with a water bound both to Na^+ and O1P.

In order to accommodate 22 *additional* water molecules, the original distribution of the 22 water molecules is fully re-arranged. Indeed for the case of 44 water

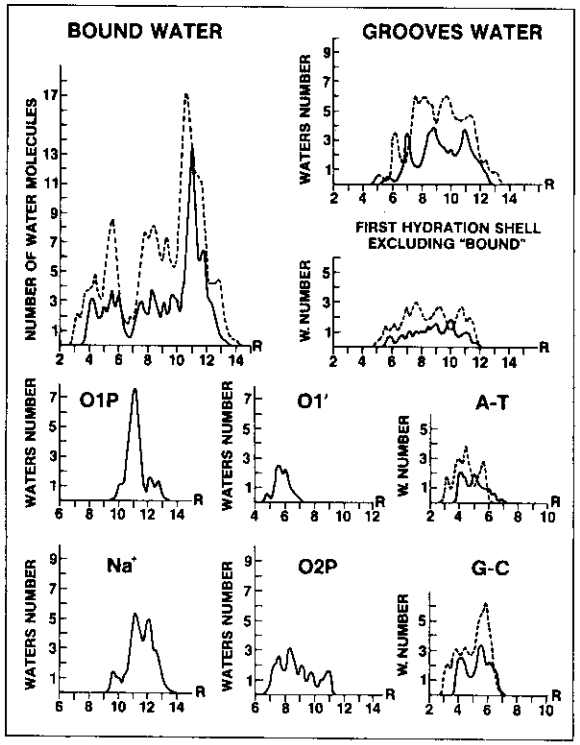


Figure 13. Probability distributions for 257 water molecules in Na^+-B-DNA.

molecules, the O3′ site starts to be interested, the O1P and the O2P site's difference, in water population, becomes smaller than for the case of 22 water molecules, and the average energy of a water molecule interacting with PO′-Na^+ drops from −155.75 Kj/mol (for 22 waters) to −148.45 Kj/mol. The rearrangement pushes the water molecules (in average) to smaller R values, and not only the hydrogen atoms of sugar but also those of the A and G bases become interested, even if marginally.

Let us now fill the DNA sample with 142 water molecules; the grooves remain essentially empty (1.43 water molecules out of 142 are found in the grooves) but the finding of some, even if very little water in it, indicates that the "bound" water number has attained its saturation value. This statement is not equivalent to saying that the structure or the maximum number of the *bound* water molecules for 142 water and for 447 water molecules solvating the DNA fragment are the same. Indeed by filling up the grooves, additional water penetrates the "bound" region, and because of the interaction between water in the "grooves" and water "bound", a rearrangement can be induced in the "bound water" structure. In other words, we are gathering preliminary evidences that the adsorption process is not "monotonic"; at low relative humidity the water molecules go mainly to the DNA sites, at higher, mainly to the grooves, at even higher to the sites and the grooves and, finally, to the "extended" grooves. Incidentally, it should be by now evident why the interpretation of experiments at different humidity (desorption-adsorption studies) are very hard (even neglecting the experimental difficulty in obtaining reproducible raw data), especially when experiments are performed on DNA fibers, where additional variables must be considered. For the case of 142 water, about 2.97, 2.18, and 1.87 water molecules are bound to Na^+, O1P and O2P, respectively; the total populations of *bound* water for PO′ and PO′-Na^+ are 4.48 and 6.40 water molecules, respectively. The total populations for the *first hydration shells* for PO′ and PO′-Na^+ are 5.71 and 6.52, respectively. The previously noted drop in the binding energy for the "bound" water molecules around PO′-Na^+ continues and the energy is now −134.48 Kj/mol. The bases are interested at the solvation and the solvation order are A<G<C<T by considering the number of "bound water" molecules, A<C<G<T by considering the number of waters in the first solvation shell and A<C<T<G by considering the binding energy for the water molecules in the first solvation shell. The distribution of the total water's population at different R values shows a well defined structure; by comparing the shift of the most intense peak one can gain a feeling of the intense water structure re-arrangement which follows the repeated addition of water molecules. The AT pair is more solvated than the GC pair for the cases of 44 and 142 water molecules, in agreement with experimental findings at low humidity. In addition we note that at low humidity O2P is more solvated than O1P, a situation that will be reversed by increasing the relative humidity. Finally, in the presence of Na^+ counter-ions, the O1P and the O2P water distributions are nearly additive, contrary to what is found for B-DNA without counter-ions.

Let us now solvate the DNA fragment with 257 water molecules (see Fig. 13). This case deserves to be analyzed more deeply, since it provides an interesting step in

the adsorption process. With 257 water molecules the first hydration shell is nearly saturated and in addition there are about 71 water molecules in the grooves. From Fig. 13 we learn that the bound water molecules exhibit the same ten peaks as found in the 447 water molecules case; however the peaks L and J (see Fig. 7) are more developed by adding water to the 257 molecules now present, because the Na^+ and the O1P sites are not fully solvated. With 257 water molecules the hydration at the base-pairs is nearly as much evolved as with 447 water molecules (peak C is, however, higher that in the case of 447 water molecules). The water molecules in the "classical" grooves are well developed at smaller R values but far from saturated at larger R values, as expected. Therefore, the term *"groove"* water could be substituted by the term *"second solvation shell"* for the case of 257 waters. Relative hydrations at the bases are not discussed, since these are given in full detail in reference 32.

In a study in progress, we are analyzing the differences in rotational freedom associated to the water molecules either bound or into the grooves at different relative humidities. This is an aspect of interest to interpret angular distribution data from scattered neutrons. We note that the findings concerning the water filaments are in agreement with the scattering data by Dahlborg and Rupprecht[27].

By considering the above partial humidity simulations, the following model for an adsorption process emerges. *First the water molecules go to the phosphate groups. If the humidity is very low* (one water molecule per phosphate—22 water case—) *then a given water has several nearly equivalent positions to select at each phosphate and the same position will not be selected. By an increase in the humidity,* such that two water molecules are available for each phosphate group (—44 waters case—), *the best solvation positions are not those left unoccupied by the first water molecule, but new ones.* This is to be expected:the Na^+-PO′ system is different from the H_2O-Na^+-PO′ system and the latter has only a vague memory of the former. By increasing the humidity such that about 6 water molecules are available for each phosphate group (142 waters case), on the average all waters solvate the Na^+-PO′ groups, but some molecules solvate even the base-pairs or are present in the groove's region. The presence of water in the groove indicates that a saturation level has been reached. Alternatively, we can state that the initial saturation of the grooves is nearly concomitant to the solvation at the base-pairs. By increasing the humidity, such that about 11 water molecules are available for each phosphate group (—257 waters case—) *more water molecules go to the groove's region and the solvation of the sugar's oxygen and of the base-pairs is nearly completed.* Finally, by making available about 20 water molecules per phosphate group (—447 water case—) additional water is packed into the groove's region and around the solvated sites, but most of it goes to the outside of Na^+-B-DNA, to the "extended grooves" region, still strongly perturbed by the Na^+-B-DNA field, so much so that the bulk water interaction is far from being reached. By adding two additional water layers (that is, by adding 1600 water molecules) one would have nearly reached the bulk water region, at an R value of about 20 A.

Conformational Transitions: A Model and a Preliminary Example

In considering the adsorption process, we should keep in mind some overall energetic aspects, reported in Fig. 14. Repeatedly, we have pointed out that the interaction of water with DNA increases by adding counter-ions to the solute therefore stabilizing Na^+-B-DNA relative to B-DNA. This increase is drastic as shown in Fig. 14, where the total (namely the sum of water-DNA and water-water) average interaction energy per water molecule is reported as a function of the number of water molecules, Nw, solvating the DNA fragment. In the top insert of Fig. 22 the full line refers to the *total* interaction, the dashed line to the *water-DNA* interaction; the *water-water* interaction is reported in the bottom insert of the figure. By a substantial increase of solvent all curves will eventually end at about —36 Kj/mol, the bulk-water value. *The total interaction could induce to think that the adsorption process is monotonic; the partial interactions offer indications for an opposite conclusion.* Physically, the following model clearly emerges *in B-DNA:* the loss of

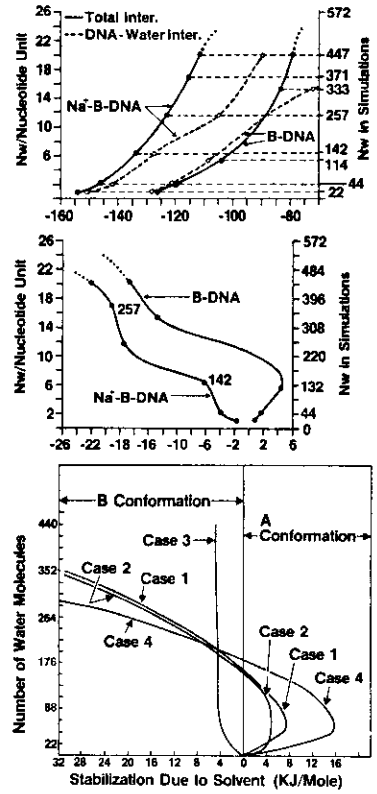


Figure 14. *Top insert:* Total interactions (solid line), water B-DNA interactions (dashed lines). *Middle insert:* water-water interaction per water molecule. *Bottom insert:* Solvent stabilizations at constant temperature (300 K) and different humidities for the A→B transition.

bulk water energy (water-water interaction) is more than compensated by the strong PO′ attraction; in addition, the water-water interaction at low humidity would amount to only a fraction of the bulk-water value. But, when the PO′ groups are screened by few solvent molecules, and when the humidity increases, then the system attempts to regain the bulk-water interaction energy since the water-water interaction becomes more and more important. *In Na^+-B-DNA* (and in any B-DNA counter-ions system) the initial adsorption situation is more complex. At low humidity, a water molecule does not experience only the PO′ field, but also the counter-ion field. The latter exerts a request on a water molecule that would demand a very different orientation from the one demanded by the PO′ field. Thus the water molecule compromises its orientation between the two conflicting requests, and by so doing an orientation is achieved, which is not as repulsive to another water molecule as the one accepted for the case of B-DNA (without counter-ions and at low humidity).

As a consequence of these effects, the water-water interaction energy presents two plateaus, where a small humidity increase brings about a strong water-water interaction variation. The two plateaus border at two steep sides, where a large humidity increase brings about only a small water-water interaction variation.

The relative humidity interval, corresponding to a variation from about 6 to about 10 water molecules per nucleotide unit, appears as a critical humidity interval. The "classical grooves" start to be filled in this interval; by a further increase in the humidity, the final saturation at the DNA solvated sites is reached and the "extended" grooves are filled. When the adsorption process is completed (very high humidity), it is hard to visualize a conformational transition in DNA (for example, from the conformation A to B), without the introduction of a third body. Equally difficult is to visualize a conformational transition at very low humidity, since the few water molecules present are fully occupied at solvating the PO′ counter-ion groups. On the other hand, in the above discussed plateau, where the incipient filling up of the classical grooves perturbs the sugar units (this is tantamount to a directional pressure) one can easily hypothesize conformational transitions. Thus combining the evidences on the ion-induced compression effect and the finding on the variations of the interaction energies with relative humidity, we can obtain a preliminary model to explain conformational transition, under specific conditions. The use of the model however, requires simulations of the type here presented for different conformations and with different counter-ions. The mathematical conditions for a conformational transition are now indicated by the following three models.

First, we consider only one type of counter-ions (for example, Na^+). Let E(A) and E(B) be the energy of DNA in conformations A and B, respectively, for DNA fragments of equal number of units. From a simulation with a equal number of water molecules and at a given temperature, one can obtain the water-DNA interaction ES(A, h(i)) and ES(B, h(j)), where the indices A and B refer to the two conformations, h(i) and h(j) to the two relative humidities (clearly different in the two conformations, since we have assumed the same number of water molecules).

A transition can occur for $\Delta E = E\Delta'$, where $\Delta E' = E(A) - E(B)$ and $\Delta E = ES(A, h(i)) - ES(B, h(j))$, assuming nearly equal entropic variations. With this work we provide one of the four needed quantities.

Second, we consider the same situation as above with, however, two different types of ions I(1) and I(2) (for example, Na^+ and Li^+) yielding either I(1)-DNA or I(2)-DNA. Let E(A, I(1)), E(A, I(2)), E(B, I(1)) and E(B, I(2)) be the energies of the fragment with either one of the two types of counter-ions and in the two conformations. Correspondingly let ES(A, I(1)), ..., ES(B, I(2)) be the corresponding simulated solvation energies as in the previous model. The introduction of more than one type of counter-ions increases the probability of finding an equality between ΔE and $\Delta E'$ since one more variable, the type of ion, has been added.

Third, let us assume the same situation as above, with, however, the additional possibility that a fractional rather than total replacement of I(i) with I(j) is considered. This last assumption considerably increases the probability of finding an equality between ΔE and $\Delta E'$. The model can be made more complete by including *a temperature index, namely by considering more than one temperature.*

As an example, let us consider the ES values, stressing however, that simulations on Li^+-B-DNA, Li^+-A-DNA, Na^+-A-DNA are not available. The different solvation energy values can be tentatively guessed by keeping in mind ionic solvation studies[35,36], i.e. the differences in the solvation energies between the A and B single helix conformation and preliminary simulation data on Li^+-B-DNA. We have made no use of the study by Marynick and Shaeffer[37] since their use of a sub-minimal basis set and the neglect of basis set super position corrections deprive their otherwise very interesting computation of quantitative validity. Their very strong attraction for a phosphate group in $(CH_3\text{-}O)_2\text{-}PO'$ with Li^+ and Na^+ (computed as -303 Kcal/mol and -198 Kcal/mol, respectively) and the very short oxygen-cation distances (computed as 1.77 A and 1.98 A, respectively) are too far from the more reasonable data obtained, however, on $(CH_3\text{-}CH_2\text{-}O)_2PO_2^-$ and yielding -152 Kcal/mol, -134 Kcal/mol, 1.93 A and 2.17 A, respectively.[30] Equivalent data, obtained from relatively old electrophoretic mobility studies, indicates that the binding to DNA for ions is $Li^+ > Na^+ > K^+$.[38]

At low humidity (1 water per nucleotide unit) Na^+-A-DNA is estimated to be more attractive to a water molecule by about 8 Kj/mol (relative to A-DNA). Simulations on A and B single helix[2,3] indicate a value between 5 Kj/mol (average value) and 17 Kj/mol (maximum difference), thus our selected value might be somewhat smaller than the correct one. At high humidity we know that the PO_4^- groups in A-DNA are sufficiently crowded, having one less water molecule strongly bound to a PO_4^- group relative to B-DNA. Thus on a sample of 447 water molecules, about 22 water molecules in A-DNA are less bound to the PO_4^- than in B-DNA by an assumed interaction of about 10 Kj/mol per water molecule. As a result the total interaction energy curve for Na^+-A-DNA crosses the total interaction energy curve of Na^+-B-DNA at a relative humidity corresponding to about 160-190 water molecules

(see Fig. 14). For Li^+-B-DNA, we have obtained preliminary data for 22 and 447 water molecules (at a temperature of 300K). At a relative humidity corresponding to 22 water molecules Li^+-B-DNA is more attractive to water than Na^+-B-DNA, but the situation is reversed at high humidity (447 water molecules) bringing about a crossing of the Li^+-B-DNA and Na^+-B-DNA total interaction energy curves at about 190-210 water molecules. Finally for Li^+-A-DNA we assume that the total interaction energy to water differs from the Na^+-A-DNA total interaction energy in the same way as found by comparing Li^+-B-DNA with Na^+-B-DNA. Until definitive simulations on Li^+-A-DNA, Li^+-B-DNA and Na^+-A-DNA will be available, the above estimates are likely all the data one can use. The stabilizations due to solvent effect for the A→B conformational transition are reported on Fig. 14 (bottom insert), (the ordinate gives the number of water molecules for either an A or a B double helix sample with twenty-two phosphate units). We consider four cases (all at 300K). In case 1, we consider the stabilization of a DNA conformation with Na^+ counter-ions, whereas in case 2, we consider the stabilization of a DNA conformation with Li^+ counter-ions. In case 1, at low humidity the form A is stabilized by water and the stabilization reaches a maximum for about 3 water molecules per nucleotide unit, then goes to zero at about 7 water per nucleotide unit, and then the B form becomes stabilized. In case 2 (Li^+ counter-ions) the same behavior is predicted, but the crossing from A to B occurs at slightly lower humidity. In cases 3 and 4, we increase not only the humidity, but we also assume that the Na^+ counter-ions of the A form are substituted with Li^+ counter-ions in the B form (case 3) or vice versa (case 4). From our preliminary data we expect that only the form B is stabilized by the solvent, whereas in case 4, the Na^+-A-DNA has a net solvation stabilization up to about 7 water molecules per nucleotide unit; at higher humidity the solvent effect helps the formation of B-DNA.

We stress that we have only referred to ΔES not to $\Delta E'$; in addition no entropic effect has been considered, or equivalently, we have assumed that the entropic contribution to the free energy is ion-independent and conformation independent at a given relative humidity and temperature. The theoretical behavior of Fig. 23, even keeping in mind its tentative nature, explains a large number of experimental findings relative to A-B transition. Clearly, the same type of reasoning can be used when considering the solvent effect of any other conformational transition. By adding to an ionic solution (containing A- or B- DNA) solvents like alcohol-water the number of water molecules available to DNA decreases because the hydrophobic part of the alcohol removes water from DNA[39,40]. Thus, if one can estimate the latter effect, then Fig. 23 provides an explanation, also for transitions, where not only the humidity and the counter-ions are varied, but also additional solvents, like alcohols, are added. Concerning the energy $\Delta E'$ for conformational transitions from A to B a value of about 84 Kj/mol has been proposed by Ivanov et al. [41] later confirmed by and Sukhorukov et al.[42].

Comment on the Second Approximation

The most crucial limitation in this approximation is in the assumption of a rigid

Na^+-B-DNA fragment, and in the position assumed for the Na^+ ions. From theoretical considerations the Na^+ lowest energy minimum (minimum a) in model compounds containing the phosphate group is in agreement with our choice. However, there is a second minimum away from the PO_2^- plane (minimum b) which will become deeper, if more than a single phosphate group is present; finally there is a third minimum along the PO bond direction (minimum c). The three minima are relatively near in energy. Let us consider the available, but indirect experimental data. A double helix chain can be constructed from the X-ray structure of sodium adenylyl (3′-5′) uridine (ApU), where one Na^+ is coordinated to the PO_4^- group and one to the two uracil carbonyl groups.[43] Thus minimum "b" is known to be a possible candidate in polynucleic acids. A double helix chain can be constructed from the X-ray structure of sodium guanylyl (3′-5′) cytidine (GpC), where the Na^+ is coordinated to the 2 free oxygen atoms of a phosphate group.[44] Thus the minimum "a" is a possible candidate. Finally, X-rays studies on crystals[45] of deoxyribose-dinucleotide sodium thymidylyl (3′-5′) thymidylate (pTpT) can be used to model a double helix with the Na^+ coordinated to one free oxygen of one PO_4 group and two oxygen atoms on two different thymine bases. This position for Na^+ can be considered as one ralated to the minumum "c". It is noted that none of the above structures refers directly to a crystal of a true polynucleic acid, nor to a 50% G-C, 50% A-T double helix structure, as in our fragment[2]. The nearest case is the one of GpC, where the Na^+ is located close to our chosen position. Most recently (when this work was completed) a large DNA fragment has been analyzed as a single crystal[47], but the counter-ions, Mg^{++}, were not detected. Several water molecules however, have been identified and assigned to the DNA fragment atoms.[60] The overall data on the water location nicely follows some of our early and above reported predictions, keeping in mind, however, the deep perturbation of one DNA fragment on the nearby fragments, the presence of impurities, and the undetermined position of the Mg^{++} ions and of the corresponding solvation water molecules[47]). Another very interesting work involving single crystals and NH_4^+ counter-ions and water molecules with a DNA fragments has recently been published.[48]

In view of the above mentioned problems concerning the determination of the counter-ion position we extended our computer experiment as reported below.

Third Approximation: "Free" Ions Simulation

The selected B-DNA double helix fragment is composed of 30 base-pairs (three full B-DNA turns). Periodically, at each turn the following sequence of ten base-pairs is selected: TA*, GC*, AT*, CG*, TA*, GC*, CG*, AT*, TA*, CG*, where the asterisk denotes the h* strand. The coordinates of this B-DNA fragment have been previously discussed.[2,31] The three B-DNA turns, see Figure 15, are hereafter referred to as *top*, *middle* and *bottom* turn, respectively.

Four-hundred water molecules and twenty Na^+ counter-ions are placed within the middle section of the cylindrical volume, shown in Figure 15. For each water

molecule we compute the interaction energy with the full three turns of the B-DNA fragment, with the water molecules and with the Na^+ ions. In addition, since our B-DNA fragment has a periodically repeated sequence for each turn, each water molecule (or each ion) is associated to two "image" water molecules (or two "image" ions) obtained by a coordinates translation (along the Z axis) of the water molecule (or ion) in the central section. In Figure 15, we have shown a water molecule in the central section of the cylindrical volume and its two "images", symmetrically located in the top and bottom sections. Equivalently, for each counter-ion in the central section, we compute the interaction with the atoms in the three DNA turns, with all the water molecules and the corresponding water images and with the other ions and ion-images. Thus we simulate a system composed by the atoms in the three B-DNA turns, by 1200 water molecules and 60 counter-ions. The solvent particles (water molecules and sodium counter-ions) in the central section are *randomly* displaced; the displacement applied to a solvent particle is also imposed on the "image" particles. Each random displacement (or "move") generates a new "configuration" for the solvent (or a new "step" in the "random walk"). An initial set of about $2x10^6$ conformations was disregarded, however, in order to erase any memory of the initial configuration for the solvent. Thereafter, the cartesian coordinates of the solvent particles (for each configuration) and the corresponding interaction energies (water with water, ions and the DNA and ion with ions and DNA) are stored on magnetic tapes, to be used later in a statistical analysis of the Monte Carlo data.

Determination of the Counter-ion Structure

In a recent and preliminary communication[49] we have reported that the sodium counter-ions in a solution with B-DNA form a pattern corresponding to two helices interwinding with the two B-DNA strands. The overall methodological approach is described at length in Ref. 29 and summarized in Ref. 50.

The most direct way to analyze the counter-ion positions is to provide the projections in the X-Y and X-Z planes (see Figure 15 for the axis choice) of the statistical distributions of the counter-ions obtained from the Monte Carlo data. The following techniques has been adopted: the entire cylindrical volume (see Figure 15) has been subdivided into small cubical cells (of 0.2 A side). A counter at each cell is activated, measuring how many times a counter-ion falls within a cell during the Monte Carlo walk. The statistical distribution is graphically visualized by reporting for each cell a number of points proportional to the number of times an ion is present in the cell. The projections of the probability distributions of the counter-ions at one B-DNA turn are reported in Figure 16. At the *bottom-left insert* we provide the X-Y projection of the ion distribution map (spots-like patterns), with an index (1 to 10) for the ten "spots" external to the phosphates and an index (11 to 20) for the ten "spots" internal to the phosphates. A counter-ion corresponds to each "spot"; the size of the "spot" provides a measure of the counter-ion mobility. The mobility of the counter-ions, determined from the probabilty distributions, is large in the x-y directions, and relatively small in the z direction; alternatively

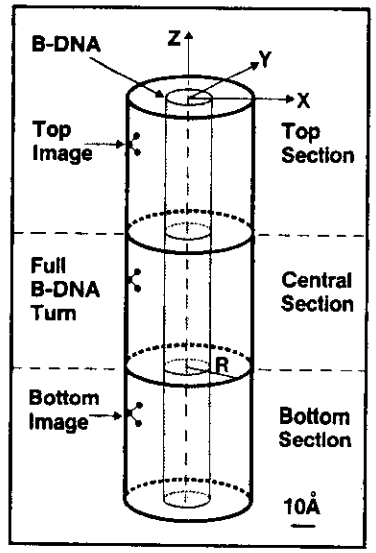


Figure 15. Volume enclosing the B-DNA fragment; "image water" molecules and "image" ions.

stated, a displacement of the counter-ions in the xy plane costs less than an equivalent displacement in the z direction (in energy terms). The phosphates of one strand are indicated by drawing the bonds between the oxygen atoms and the corresponding phosphorous atom and by explicitly indicating the corresponding bases, A, G, C and T. The atoms of the phosphates in the second strand are indicated simply with dots at the nuclear positions and for the corresponding bases we report only a dot for the nitrogen atom position (the one connecting the base with the sugar). *The projections of the twenty counter-ion probabilities form two nearly regular circular patterns.* In the bottom-right insert of Figure 16 we present the same probability density distributions, this time projected into the X-Z plane. In this insert the phosphate groups of the two strands are enclosed into helical envelopes and the base-pairs are identified by reporting the base-pair molecular plane. In the top-right insert we repeat the X-Z projection, this time connecting into an envelope those counter-ion probability distributions which are nearest neighbors; to simplify the diagram we have not enveloped the phosphate groups into two helices as done in the bottom right insert. *The pattern emerging from these three inserts is repeated in the top-left insert, where we draw the two helical envelopes for the phosphates of the two strands, and the two helical envelopes for the sodium counter-ions, one penetrating into the major groove and the other outside to the minor groove.* The counter-ions helices are designated by the letters H and H*. From the density projections it is evident that the H helix is external to the cylindrical volume determined by the phosphate groups, whereas the H* helix is internal to it. The cross section of the imaginary "cable" enveloping the ions of H* is larger than the corresponding crossection of the "cable" enveloping the ions of H. The

physical reason is rather obvious:the counter-ions in H* are strongly affected by the base-pairs. Therefore the *exact position* for an ion in H* *is base-pair sequence dependent; the* "irregularities" of the distributions 11 to 20 in the bottom-left insert of Figure 25, are the effect of the base-pair sequence dependence.

The new structure, H and H, obtained in our computer-experiment is physically very pleasant since it optimizes at once several basic energetic requirements:* 1) it keeps the counter-ions as far as possible away one from the other but at the same time; 2) it satisfies the *very strong attraction* between the Na^+ ions and the free oxygen atoms of the phosphates as well as 3) the *strong attraction* between the Na^+ ions and the bound oxygen atoms in the phosphate groups and the *attraction* to the base-pairs, 4) it allows *both* the phosphate groups and the counter-ions to be solvated by the water molecules, thus making use of the solvation energy to stabilize the entire system.

Biologically this structure is very interesting since:1) *it provides for a base-pair recognition mechanism at long range distances due to the base-pair sequence dependent H* structure, 2) it can easily allow for the exchange of one or more sodium counter-ions with different counter-ions (from Na^+ to K^+ or Mg^{++} or Ca^{++} etc. etc.) 3) it can account for rapid structural and conformational reorganization processes, typical of ionic solution with macromolecules as solute and 4) because of the ionic mobility and very strong interactions, it can act either as an important sink (or source) of energy.*

Determination of the Water-Structure Solvating DNA

The statistically most probable distribution of the water molecules (at 300 K temperature) either *bound* or in *the first solvation shell* for an atom (or groups of atoms) in B-DNA is analyzed in Figures 17 to 21. In these figures the probability distribution for the hydrogen atoms is given as a *dotted line* and the one for the oxygen atoms is given as a *full line.* On one axis (ordinates) we report the number of hydrogen or oxygen atoms as a function of R (given in the abscissa, in A units); we recall that R is the distance of the hydrogen (or oxygen) atoms from the Z axis (see Figure 15). The notations B and FS differentiate between *bound* and *first solvation shell* water molecules; the notations h and h* differentiate the two strands.

In Figure 17, we report the analysis for water molecules *bound* to the free oxygen atoms (O1P and O2P), and to the bound oxygen atoms (O5′ and O3′) in the two strands (h and h*). These analyses of water molecules bound at atomic sites are complemented with the *bound* water distributions and the *first solvation shell distributions* at the PO_4^--CH_2 group.

We learn that *at the O1P* sites the water oxygen atoms in the h strand have a maximum at about 11 A, whereas the maximum is at 12 A in the h* strand; the water hydrogen atom distributions (dotted lines) in h differ from those in h*, therefore indicating different orientational arrangements at the two strands. The

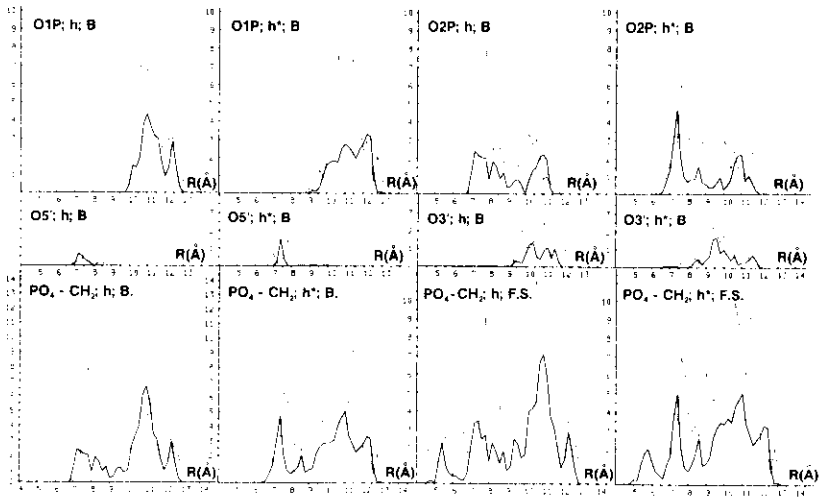


Figure 17. *Top* and *Central* inserts. Distribution of the hydrogen and oxygen atoms of water molecules bound to the oxygen atoms in the PO_4^- groups; *bottom inserts:* equivalent distribution for the PO_4^--CH_2 group either considering bound (first two insert from the right) or first solvation shell water molecules (last two inserts from the right).

same holds for the water molecules at O2P. There is little, if any, water at the O5′ as previously noted and not much water at O3′. This information is iterated by providing the distribution for the water molecules bound at the PO_4^--CH_2 site; as discussed above the first solvation shell can extend much further than the *bound* water distribution, as clearly seen in the figure. The integral of the distribution values of the hydrogen and oxygen atoms as function of R provides the number of water molecules *bound* to the atoms or groups of atoms, reported in this figure. These are given in Table V, both for *bound* and *first hydration shell* water molecules. In the following, we shall (in general) not comment on the features which can be obtained by inspection of the self-explanatory figures. We feel that infrared, Raman, nmr, neutron, X-ray experiments on DNA in solution will now be more easily interpreted, such data being available.

In Figure 18, we report the analysis for the water molecules *bound* to the lone pair nitrogen (N3 and N7) and oxygen (O2, O4 and O6) atoms of the bases. The analysis of the water molecules bound to atoms belonging to the bases is extended in Figure 19 when we report the distributions for the water molecules bound to specific NH_2 groups and the *average* distributions for water molecules bound to NH_2 or oxygen or nitrogen atoms at the bases obtained by considering all the bases at the two strands.

In Figure 20, we compare the distributions of water molecules bound at the O1′ (of

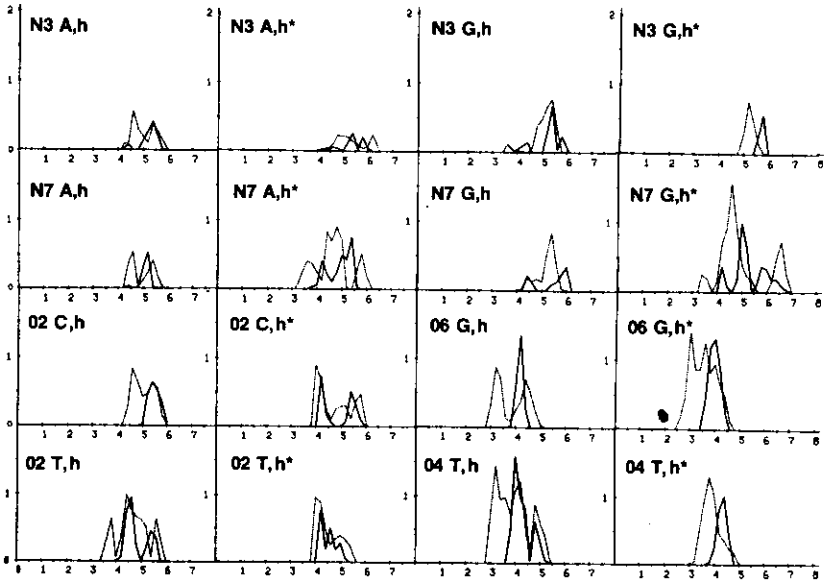


Figure 18. Distribution of hydrogen and oxygen atoms for water molecules *bound* at the nitrogen (N3 and N7) or oxygen (O2, O4 and O6) of bases forming either the h or the h* strand in B-DNA.

sugar) or the first solvation shell of the entire sugar unit. In this figure we also report the distribution of the water bound to the bases (in each strand).

In Figure 21, we complete the analysis of the bases, by reporting the average distribution of the water molecules bound to the A-T and G-C base-pairs, and the average distribution for the base-pairs either in the h or h* strands, and the first hydration shell distribution at the bases (average values for both strands).

These very detailed but coincided graphical presentations of the solvation in B-DNA, are complemented by the data of Table V.

The analysis above reported for 400 water molecules (per B-DNA turn) has been extended to intermediate and low relative humidity; we have now concluded our computer experiment by considering 380, 240, 220, 180, 140, 40, and 20 water molecules per B-DNA turn. *The H and H* structures for the counter-ions have been found at each relative humidity; therefore it is very reasonable to assume that such structures are present in solution.*

Determination of the Water Structure Solvating the Counter-ions

In Figure 22, we report the distribution of the water molecules bound to the ten

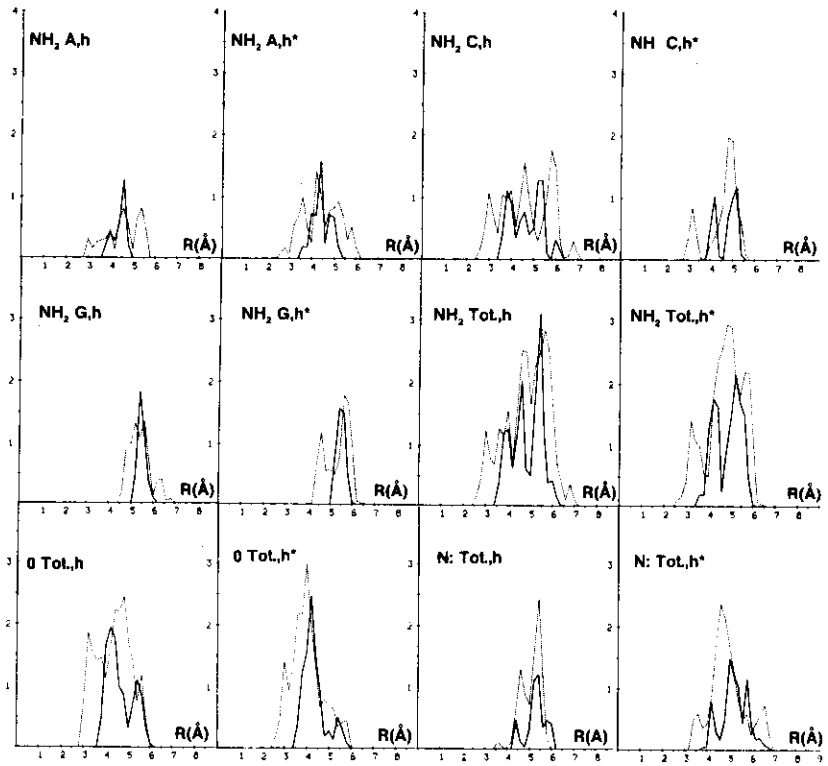


Figure 19. *Top and central* insert: distribution of hydrogen of oxygen atoms for water molecules *bound* at the NH_2 groups of the bases. The four bottom inserts report equivalent quantities but for an average oxygen (either O2, O4 or O6) and an average nitrogen (either N3 or N7) in the bases at the two anti-parallel strands, h and h*.

counter-ions in the H helix (ions 1 to 10) and to the ten counter-ions in the H* helix (ions 11 to 20). In this figure the counter-ion is placed at the *origin* of the axis. The orientation of the water molecule (oxygen nearer, hydrogen farther away) is very typical of a sodium ion in solution[13,35,36]. Each counter-ion is solvated; this finding is expected to be valid also for K^+ counter-ions, and most likely for the Li^+ counter-ions.

In Figure 23, we report the distribution for the water molecules in the two strands either *bound* or in the first hydration shell. From these data we see clearly that the two strands have a different hydration pattern, as also reported in Table V. This very important finding remains hidden when one reports the distributions for the total system of 400 water molecules solvating the ten base-pair turn in B-DNA or when one considers the water molecules in either the first solvation shell or those

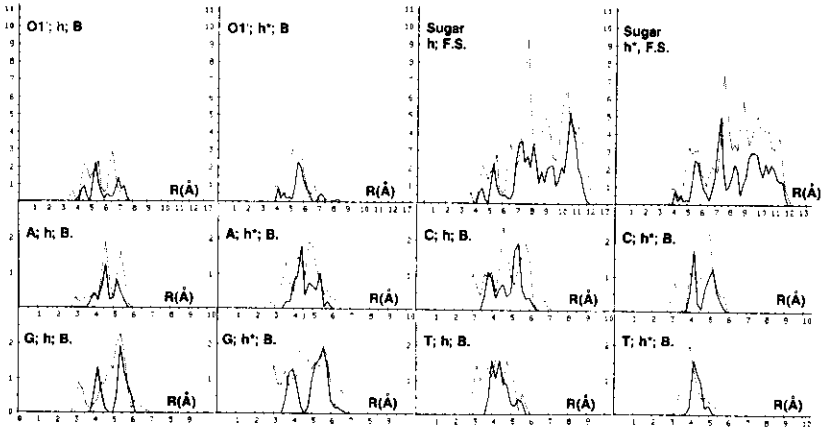


Figure 20. *Top inserts:* water characterization at the sugar units in the two strands:bound water at O1′ atom, and *first solvation shell* water in the sugar units. *Middle and Bottom insert:* distribution for water molecules *bound* to the bases.

bound to B-DNA. In Figure 24, we present the distribution of water molecules bound to B-DNA and to the ions, those in the first hydration shell and those in the grooves.

In the relative humidity range from 400 to 240 water molecules (per B-DNA turn) we find all the time about *four water molecules* per Na^+ counter-ion. By decreasing the relative humidity, namely for the cases of 220, 180, 40 and 20 water molecules (per B-DNA turn) the *average hydration number* for the 20 counter-ions decreases to 3.8, 3.5, 1.5 and 0.8, respectively. Therefore, at low humidity a water molecule solvates either the counter-ions, or B-DNA. This result should warn against extrapolations (relating to B-DNA in solution) of quantum mechanical computations obtained by considering few water molecules and one counter-ion.

In Figure 25, we report the average energy for a water molecule in the volume R and R+dR; the energy is decomposed as water-water, water-DNA, water-ions. Notice how this energy is nearly constant from small to large values of R.

Base-Pair Recognition

The important conclusion from our computational experiment is that the counter-ions and the solvation water molecules form two different patterns at the two DNA strands. We designate the *global system* composed by the h strand, the counter-ions in the H helix and the water molecules bound to h and H as the S *"super-strand"*. The equivalent global system for h* and H* and the solvating water molecules is designated as *S**. As known, the two strands h and h* differ only because they are anti-parallel. This difference, however, is enhanced by the counter-ions distribution and by the water molecules. Any biological process dealing with DNA in water

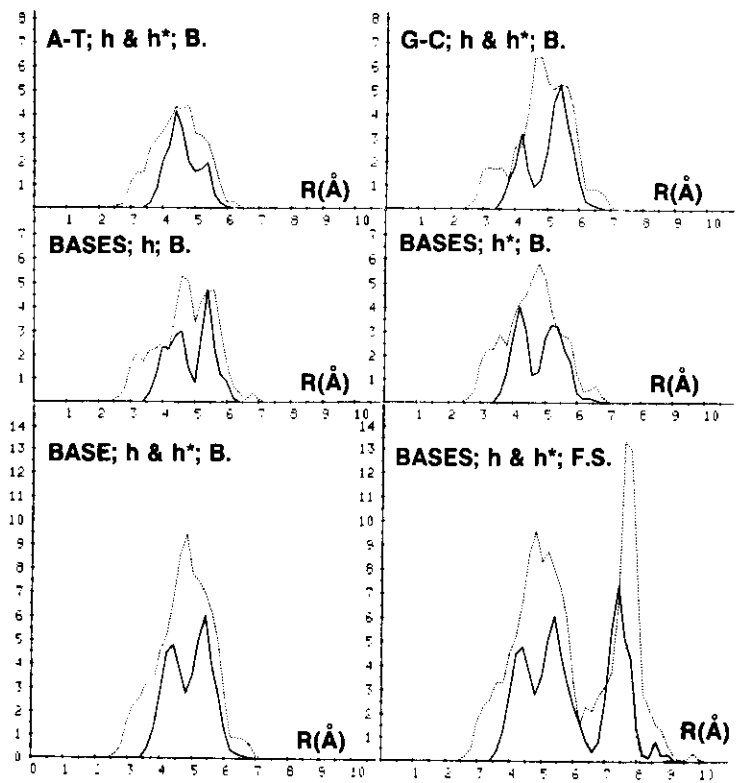


Figure 21. *Top inserts:* Distribution for water molecules at the base-pairs. *Middle inserts:* average distribution for water molecules *bound* in the two strands; *bottom:* average distribution for both strands of water molecules *bound* (left) and in the first solvation shell (right) of the bases.

solution deals with the S and S* super-strands and not only with the h and h* strands. In addition, to a given DNA conformation, there is a corresponding wide spectrum of S and S* conformations. Indeed, the structure of H and H* is dependent upon the *counter-ion charge*, the *ionic radius* and the *counter-ion concentration* for a given *temperature*.

Let us consider a few immediate implications of the above findings. In the following we present a *"base-pair sequence"* recognition mechanism. *Let us consider a molecule*, for example a glycine zwitterion, *approaching DNA in water solution, but still relatively far away from DNA, such that the direct interaction glycine-base-pairs can be assumed as small.* For example, we assume that the C^{α} of glycine is at 17 A on the x axis (with y=0; see Figure 15) and optimally oriented relative to our B-DNA fragment, the 1200 water molecules and the 60 sodium counter-ions. Fur-

Table V
Number of water molecules and its interaction energy (in Kcal/mol) with atoms or groups of atoms of B-DNA and the Na^+ ions.

Atoms or Groups	Number of Water Bound		Average Energy	
	h	h*	h	h*
O1P	3.24	3.36	−37.97	−34.57
O2P	2.92	3.05	−38.33	−41.11
O5′	0.18	0.23	−37.76	−43.52
O3′	0.89	1.12	−37.51	−34.24
O1′	1.19	1.13	−39.00	−37.81
N3 in A	0.49	0.26	−39.07	−34.97
N7 in A	0.46	0.94	−37.97	−37.73
N3 in G	0.72	0.29	−36.70	−39.74
N7 in G	0.60	1.20	−33.03	−41.94
O2 in C	0.67	0.97	−39.91	−37.33
O6 in G	1.08	1.34	−35.91	−40.53
O2 in T	1.00	1.00	−41.08	−38.94
O4 in T	1.62	1.26	−36.97	−29.29
Na^+	4.00	4.40	−39.72	
PO_4^--CH_2	6.43	7.18	−38.04	−37.09
NH_3 in A	1.49	1.77	−39.79	−37.65
NH_3 in C	2.76	2.31	−38.24	−39.18
NH_2 in G	1.94	1.46	−36.23	−39.93
NH_2 in A, G, C	2.13	1.70	−37.81	−39.04
N in A, G	0.57	0.67	−36.40	−38.85
O in C, G, T	1.00	1.06	−38.61	−38.35
A	2.42	2.55	−40.03	−37.46
C	3.43	3.29	−38.58	−38.43
G	3.84	4.15	−35.77	−40.68
T	2.62	2.26	−38.51	−37.68
Grooves	131.82		−30.41	

ther, we assume that glycine is *translated* along the z axis by steps of 0.25 A, re-optimizing its *orientation* at each value of Z.

The interaction of the counter-ions with glycine is rather large (long range interaction of ionic nature), and the interaction of the H* counter-ions with the base-pair has been previously shown to be large and base-pair sequence dependent. Therefore, glycine will recognize the base-pair sequence, via the counter-ions. *The proposed recognition mechanism is a relay-type mechanism: base-pairs to counter-ions, counter-ions to glycine.* A disordered pattern in the ions, rather than the ordered one we have determined, will lead to no base-pair recognition. Only a *special* ordered pattern lead to recognition. Further, the recognition in our model is dependent upon *ion concentration,* ionic *radius* and temperature. This latter comment is of importance in study of the evolution of genetic proto-materials. Among feasible application of this proposed mechanism we mention: 1) recognition of a sequence perturbed by cancer or an anti-cancer intercalating molecule and 2) recognition of amino acids by RNA in protein syntheses.

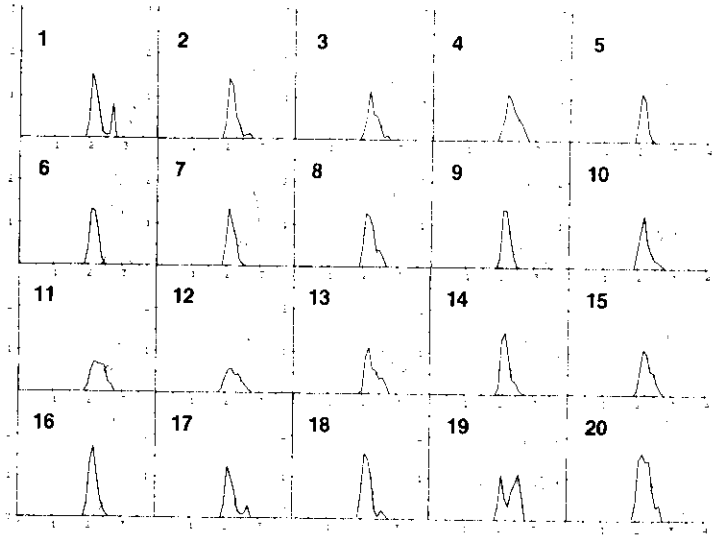


Figure 22. Hydrogen and oxygen atoms distributions for water molecules *bound* to the counter-ions in the H $helix^{1-10}$ and in the H* $helix^{11-20}$.

In Figure 26 (left insert), we report the interaction energies of GLY (in Kcal/mol) with the atoms of the B-DNA fragment, with the water molecules solvating B-DNA with the counter-ions and the total interaction energy. The interaction of GLY with water is nearly constant and repulsive (screening effect). The interaction of GLY with B-DNA shows a *low* frequency periodicity associated with the B-DNA turn and a high frequency periodicity associated with the nucleotide units. The interaction with the counter-ions shows the low frequency periodicity but with opposite phase as the one for the GLY-B-DNA interaction. The GLY counter-ion interaction is attractive and over compensate (being larger) the repulsive interactions with water and B-DNA. The high frequency spikes in the GLY-B-DNA and in GLY counter-ions are separated by about the same distance as the base-pair to base-pair distance. Clearly the pattern will differ in A-DNA, and it will *locally* differ if a molecule intercalates DNA. Notice that the "recognition spikes" in the total interaction energy are about 10 Kcal/mol, namely a value sufficiently large for being very important in biological mechanisms, but also sufficiently low as to be affected by thermal effects. Notice how one pair has a pattern different from another pair. To our knowledge this figure represents the first quantitative energetic representation of the reading of DNA by a molecule (GLY in our experiment).

DNA Unwinding

Another implication of our findings concerns the unwinding mechanism in the double helix. As known, a double helix structure has a critical temperature and a

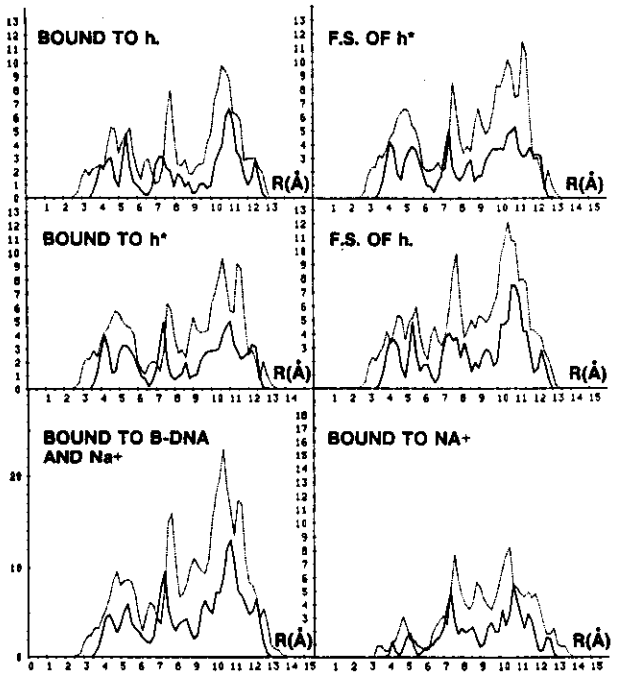


Figure 23. *Top and central* inserts: average hydrogen and oxygen atoms distributions for water molecules in the h and h* B-DNA strands either *bound* or in the *first hydration shells. Bottom inserts:* water molecules bound to B-DNA and to the counter-ions (left) or only to the counter-ions (right).

critical ionic concentration, just beyond of which the two helices extremely rapidly snap apart cooperatively. By a 20 K increase in the temperature (in our simulated system) we obtain a different counter-ions pattern with the counter-ions in the H* helix closer to the base-pairs than at 300 K. We recall that the interaction of an Na^+ ion with the base-pairs, is not only strong when Na^+ is at the perimeter of the base-pair (in the plane containing the base-pair skeleton) but also when Na^+ is *above* the base-pair. In this position the attraction "base-pair to Na^+" is opposed by the hydrogen atoms forming the base-pairs hydrogen bonds. An increase in the system thermal motion (due to temperature) can bring about a separation between two successive base-pair and/or hydrogen bonds breakage within a base-pair. In either case, a sodium ion can approach the bases even further, and oppose the restoration of the original DNA configuration.

The disruption of the double helix structure following progressive removal of counter-ions at constant temperature is easily understood in terms of the large stabilization brought about by the ions to the system "water and DNA"[32]. We note

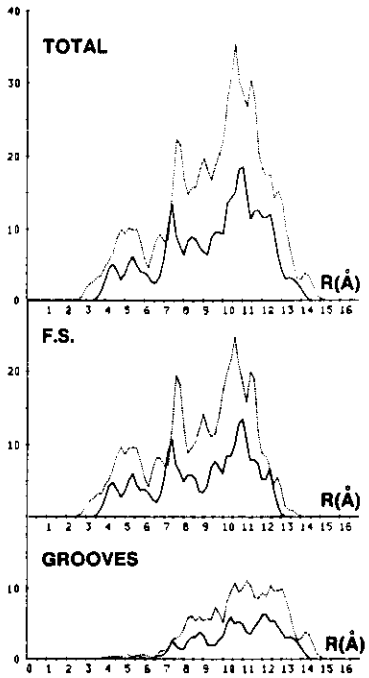


Figure 24. Statistical hydrogen and oxygen atom *total* distribution for 400 water molecules solvating one B-DNA turn and 20 Na^+ counter-ions (top) in B-DNA *first solvation shell* (middle) and in the *grooves* (bottom).

that the total energy of the system reported in this work is more stable by about 20 Kj/mol, than the system analyzed in Reference 32.

From our energy data we estimate that the DNA and the counter-ions field extends up to about R=25 A. Therefore, we expect that X-ray crystal studies from single crystals should show evidences of the DNA to DNA perturbation. As a consequence the counter-ion structure in a single crystal is expected to differ from the counter-ion structure of DNA in solution.

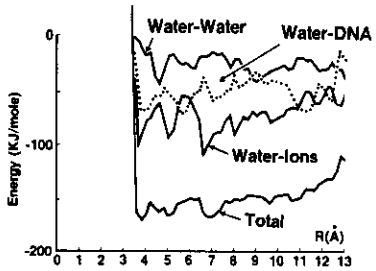


Figure 25. Average interaction energy (Kj/mol) for water with the remaining water molecules (water-water), with the B-DNA fragment (water-DNA), with the ions (water-ions) and the total interaction as function of R.

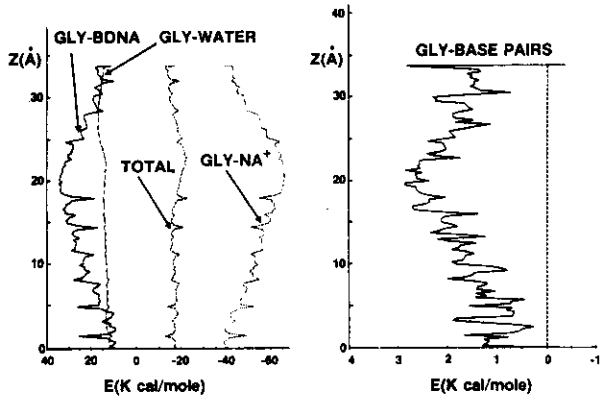


Figure 26. Recognition by GLY of DNA:Interaction energies (in Kcal/mol) of GLY with B-DNA, with the water molecules solvating DNA, with the counter-ions and total interaction; interaction energy of GLY with base-pairs is shown on the right-hand side panel.

Conclusions

Study is in progress for the determination of the counter-ion structure for Li^+, K^+, Mg^{++} and Ca^{++}. In addition, the approximation of the rigid solute is under scrutiny.[51,52] In a recent work[53] on the agar-agar double helix, the rigid structure in the solute has been partially relaxed; the experience gained in this recent work will be transferred to future DNA simulation. These computer experiments are now being extended in order to assess the optimal number of counter-ions for our B-DNA fragment. *From preliminary simulations, at low relative humidity and 300 K,* we obtain that 19 sodium counter-ions (per B-DNA turn) bring about a *net stabilization* in the total interaction energy of about 0.58% relative to the case with 20 sodium-ion (per B-DNA turn). This net stabilization results mainly from about a 5% decrease in the ion-ion repulsion and an increase of about 0.10% in the ion-DNA attraction. If we simulate (at the same low relative humidity and at 300 K, as above) 21 counter-ions per B-DNA turn, then we obtain a net *destabilization* of about 0.64% in the total interaction energy of the system, relative to the case with 20 sodium counter-ions per B-DNA turn; the destabilization is mainly the effect of an increase of the ion-ion repulsion (by about 5%) and a decrease (0.07%) in the ion-DNA attraction. When we consider either 18 or 17 or 10 counter-ions, we obtain no additional stabilization for the 18 ions case (relative to the 19 ions case) and we notice a destabilization for the 17 and 10 ions cases.[54] This type of study should allow us to combine our micro-analyses with the thermodynamical models presented by Record et al.[55] and by Manning.[56]

The model proposed for conformational transitions (see the section Conformational Transitions) can be adapted most easily, from the simulations where the counter-ions are fixed at a predetermined position, to simulations where the counter-ions

are mobile in the water solvent. The specific solvation energy of different ions in water, corrected by the specific DNA field effect (H and H* structures) constitutes an appreciable and ion-specific contribution to ES(A) and ES(B). Computer experiments in progress[54] appear to nicely reproduce laboratory ion-induced conformational transitions[39-42,57] and observed melting point at different ionic strength.[58]

In this review, we have omitted a detailed analysis of the approximations adopted, and we have not stressed the generality of our approach for the *"computer experiments"*, reported. The interested reader can find such information in references 29 and 50.

One trend is becoming more and more evident:the very approximated nature of the "quantum-mechanical rationalizations" of laboratory experiments, (rather conspicuous in the sixties and still retained, for example, in studies based on approximated electrostatic potentials) is becoming more and more apparent[59] and therefore, there is an increasing reliance on those theoretical formulations, where essential parameters like *temperature, statistical distributions, time, solvents* and *reaction fields* are no longer ignored. Indeed these formulations, proposed about twenty to thirty years ago and generally, long neglected in quantum-biophysics, are complementary to laboratory experiments. This trend is emerging not accidentally, but because more and more attention is given to dynamical aspects in nucleic acid and protein chemistry.

References and Footnotes

1. Scordamaglia, R., Cavallone, F. and Clementi, E., *J. Am. Chem. Soc. 99,* 5545 (1977).
2. Clementi, E. and Corongiu, G., *J. Chem. Phys. 72,* 2979 (1980).
3. Clementi, E. and Corongiu, G., *Biopolymers 18,* 2431 (1979).
4. Clementi, E. and Corongiu, G., *Int. J. Quant. Chem. 116,* 897, (1979).
5. Clementi, E. and Corongiu, G., *Chem. Phys. Letters 60,* 175 (1979).
6. Clementi, E. and Corongiu, G., *Gazz. Chim. It. 109,* 201 (1979).
7. Romano, S. and Clementi, E., *Int. J. Quant. Chem. 17,* 1007 (1980).
8. Clementi, E., Corongiu, G., Jonsson, B. and Romano, S., *FEBS 100,* 313 (1979).
9. Clementi, E., Corongiu, G., Johnsson, and Romano, S., *J. Chem. Phys.* 260 (1980).
10. Lewin, S., *J. Theor. Biol. 17,* 181 (1967).
11. Clementi, E., and Corongiu, G., *Biopolymers 20,* 551 (1981).
12. Lie, G. C., Yoshimine, M. and Clementi, E., *J. Chem. Phys. 64,* 2314 (1976); Matsuoka, O., Yoshimine, M. and Clementi, E., *J. Chem. Phys. 64,* 1351 (1976).
13. Clementi, E., *Determination of Liquid Water Structure,* Lecture Notes in Chemistry, Vol. 2, Springer-Verlag, Berlin, (1976).
14. Falk, M., Hartman, K. A. and Lord, R. C., *J. Am. Chem. Soc. 84,* 3843 (1962).
15. Falk, M., Hartman, K. A. and Lord, R. C., *J. Am. Chem. Soc. 85,* 387 (1963).
16. Falk, M., Hartman, K. A. and Lord, R. C., *J. Am. Chem. Soc. 85,* 391 (1963).
17. Rupprecht, A. and Forslind, B., *Biochim. Biophys. Acta 204,* 304 (1970).
18. Hearst, J. E. and Vinograd, J., *Proc. Natn. Acad. Sci. U. S. A. 47,* 825 (1961).
19. Hearst, J. E. and Vinograd, J., *Proc. Natn. Acad. Sci. U. S. A. 47,* 999 (1961).
20. Hearst, J. E. and Vinograd, J., *Proc. Natn. Acad. Sci. U. S. A. 47,* 1005 (1961).
21. Wolf, B. and Hanlon, S., *Biochemistry 14,* 1661 (1975).
22. Tunis, M. J. B. and Hearst, J. E., *Biopolymers 6,* 1325 (1968).
23. Tunis, M. J. B. and Hearst, J. E., *Biopolymers 6,* 1345 (1968).

24. Kuntz, I. E., Branfield, T. S., Law, G. A. and Purcell, G. V., *Science 163,* 1329 (1969).
25. Privalov, P. L., Ptitsyn, O. B. and Birshtein, T. M., *Biopolymers 8,* 559 (1969).
26. Texter, J., *Prog. Biophys. Molec. Biol. 33,* 83 (1978).
27. Dahlborg, U. and Rupprecht, A., *Biopolymers 10,* 849 (1971).
28. Corongiu, G. and Clementi, E., *Gazz. Chim. It. 108,* 687 (1978); *108,* 273 (1978).
29. Clementi, E., Lecture Notes in Chemistry Vol. 19, *Computational Aspects for Large Chemical Systems* Springer-Verlag, Heidelberg, Berlin, New York, (1980).
30. Clementi and E., Corongiu, G., & Lelj, F., *J. Chem. Phys. 70,* 3726 (1979), (and references, thereby given). For the Na^+ and Li^+, K^+, Mg^{++}, Ca^{++}, atom-atom pair potentials with DNA see G. Corongiu and E. Clementi (to be published).
31. Fieldman, R., *Atlas of Macromolecules* document 13 (1976), Natl. Inst. Health, Bethesda, Maryland, U. S. A. For earlier references see: Arnott, S. and Hukins, D. W. L., J. Molec. Biol., *81,* 93 (1973). Arnott, S. and Hukins, D. W. L., Biochim. Biophys. Res. Commun., *47,* 1502 (1972).
32. Corongiu, G. and Clementi, E., *Biopolymers* (in press).
33. Perahia, M. S. J. and Pullman, B., *Biochim. Biophys. Acta 474,* 349 (1977), and references therein given. The very extended number of papers referenced in 33 are very similar in the technique adopted.
34. Ranghino, G., and Clementi, E., *Gazz. Chim. It. 109,* 170 (1978).
35. Barsotti, R. and Clementi, E., *Theor. Chim. Acta 42,* 101 (1977). See also Marynick, D. A. and Schaffer, III, H. F., *Proc. Nat. Acad. Sci. U. S. A.* 3794 (1975).
36. Clementi, E. and Barsotti, R., *Chem. Phys. Letters 59,* 21 (1978).
37. Marynick, D. S. and Schaeffer, H. F., *Proc. Nat. Acad. Sci. U. S. A. 72,* 3794 (1975).
38. Ross, P. D. and Scruggs, R. L., *Bioplymers 2,* 79 (1964).
39. Frisman, E. V., Slonitsky, S. V. and Vaselkov, A. N., *Int. J. Quant. Chem. 16,* 847 (1979).
40. Frisman, E. V., Vaselkov, A. N., Solnitsky, S. V., Karavaev, L. S. and Vorob'ev, V. E., *Biopolymers 13,* 2169 (1974).
41. Ivanov, V. I., Zhurkin, V. B., Zavriev, S. K., Lysov, Yu. P., Minchenkova, L. E., Minyat, E. E., Frank-Kametskii, N. D. and Schyolkina, A. K., *Int. J. Quant. Chem 16,* 189 (1979), see in addition Zhurkin, V. B., Lysov, Yu. P. and Ivanov, V. I., *Biopolymers 17,* 377 (1978).
42. Sukhorukov, B. I., Gukowsky, I. Ya., Pekrov, A. I., Gukowskaya, A. S., Mayevsky, A. A. and Guenkova, N. M., *Int. J. Quant. Chem. 17,* 339 (1980); Bunville, L. G., Geiduschek, E. P., Rawitscher, M. A. and Sturdevant, J. M., *Biopolymers, 3,* 213 (1965).
43. Seeman, N. C., Rosenberg, J. M., Suddath, F. L., Kim, J. J. P. and Rich, A., *J. Mol. Biol. 104,* 109 (1976).
44. Rosenberg, J. M., Seeman, N. C., Day, R. O. and Rich, A., *J. Mol. Biol. 104,* 145 (1976).
45. Camerman, N., Fawcett, J. K. and Camerman, A., *J. Mol. Biol. 107,* 601 (1976).
46. Drew, H. R., Takano, T., Tanaka, S., Ikatura, K. and Dickerson, R. E., *Nature 286,* 567 (1980).
47. Drew, H. R. and Dickerson, R. E., *J. Mol. Biol* (in press). We thank the above authors for having sent us a preprint of this paper.
48. Klug, A., Jack, A., Viswamitra, M. A., Kennard, O., Shakked, Z. and Steitz, T. A., (1979), *J. Mol. Boli. 131,* 669 (1979), and references therein given. We thank Dr. Kennard for preprints of their work.
49. Clementi, E. and Corongiu, G., *Int. J. Quant. Chem.* (submitted) (1981).
50. Clementi, E., IBM *J. Res. and Dev. 25,* 315 (1981).
51. Schellman, J. A., *Biopolymers, 13* 217 (1974).
52. Olson, W. K., *Nucleic Acid Geometry and Dynamics* Sarma, R. H., Ed., Pergamon Press, New York, pg. 383 (1980).
53. Fornili, S., Corongiu, G., Palma, U. and Clementi, E., (to be published).
54. E. Clementi (to be published).
55. Record, M. T., Jr., Anderson, F. C. and Lohman, T. M., *Quart. Rev. Biophys. II, 2,* 103 (1978); Anderson, C. F. and Record, M. T., *Biophys. Chem., 77,* 353 (1980).
56. Manning, G.S., *Quart. Rev. Biophys. II,* 2179 (1978); see also Gueron, M. and Weinsbuck, G. *Biopolymers 19,* 353 (1980).
57. Chan, A., Kilkuskie, R. and Hanlon, S., *Biochem., 18,* 84 (1979); Anderson, P. and Bauer, W.

Biochem., 17, 594 (1978); Pohl, F. and Jovin, T., *J. Mol. Biol. 67,* 375 (1972); Hanlon, S., Chan, A. and Berman, G. *Biochem. Biophys. Acta, 519,* 526 (1978).
58. Schildkraut, C. and Lifson, S. *Biopolymers, 3,* 195 (1965).
59. Clementi, E., *J. Phys. Chem., 84,* 2122 (1980).
60. Dickerson, R. E., Drew, H. R., and Conner, B., in *Biomolecular Stereodynamics, Volume I,* Ed., Sarma, R. H., Adenine Press, NY p. 1-xx (1981).
61. Work partially supported by the National Foundation for Cancer Research.

Structure of Water and Counterions for Nucleic Acids in Solution

Enrico Clementi[6]
IBM Corporation, IS/TG
Poughkeepsie, New York 12602

Introduction

We shall review several statistical-mechanical computer experiments, performed to elucidate concepts needed to discuss water molecules hydrating and/or solvating biomolecules. In addition we shall use such concepts and methods to determine the positions and orientations of water molecules interacting with nucleic acids. Since DNA's helices are unstable without counterions (in the solvent) we have considered Li^+, Na^+ and K^+ as counterions. We recall that among these mono-charged counterions Na^+ and K^+ are biologically very important. As known, the determination of their position by X-ray diffraction or other techniques is still problematic. Recently, however, a mono-charged counterion, NH_4^+ and the waters of crystallization in a single-crystal of a small DNA fragment have been analyzed[1] leading to a tabulation for the positions of a large number of water molecules.

It should be noted that X-ray studies of hydrating structures have also been reported for other single-crystals containing divalent-ions and several DNA small fragments.[2-5] We shall not, however, refer extensively to these studies, since the determination of the water molecule's position and orientation appears to be still in progress and no final and reliable tabulation of the water molecule positions has been feasible. Often, in these papers, the water molecule orientations are guessed, in order to provide with proposals (on the hydrogen bonded organization) in agreement with available spectroscopical and physico-chemical experiments and more recently with predictions from computer experiments. Therefore the diffraction data on DNA's hydration provide an upper and a lower-bound to hypotheses and predictions, with a very broad range, however, since obtained at low resolution and, presently, rather unrefined. In general, extrapolations from studies on DNA's olygomers to problems related to DNA's polymers are still limited by the absence of reliable studies: 1) on the head-tail effects as function of the chain length, 2) on the effects introduced by the presence of spermidine and other "impurities", 3) on the quantitative effects due to interactions and cooperative forces between olygomeric units, and 4) on the low resolution, which brings about well known problems in particular for the determination of the position and orientation of water molecules and the counterions. These studies[1-5] on the DNA single-crystals are nevertheless very important, since they have provided direct experimental evidences on the structure of old and of some totally new DNA conformers; as known, previously only diffraction data from fibers were available.

From the above, one can correctly conclude that notable progress has been made in the last few years concerning the determination of the *position* of water molecules at DNA. There are now reasons to hope for an equivalent progress in the determination of the *orientation* of the hydrating structures of water and in the location of the position for mono-charged counterions, like Na^+ and K^+. Neutron beams and higher resolution X-ray diffraction data, as well as NMR and computer experiments, are expected to yield the above improvements.

In this work we shall first recall few concepts which, generally, are presented in papers dealing with the theory and methodology of solvent-solute interactions; this is done in order to introduce new methods and concepts, some of which are presented in this work for the first time. It is noted that these concepts and techniques can be generalized, leading to notably more powerful simulation techniques than those today available. After presenting this introductory material, we shall review recent computer experiments performed at our laboratory on the hydration of nucleic acids.

A monography[6] and some of the papers in this volume can be used as an introduction and/or extension of this work. Our methodological approach is a *"generalized model"* consisting of an ordered set of sub-models (or levels), labeled 1...,i,...,N linked by the postulate that the *output from the sub-model "i" is the input to the sub-model "i+1"*. In general, the particles in the sub-model "i" are the components of the particles in the sub-model "i+1"; as a consequence, the equations of motion (and the underlaying statistic) evolve from sub-model to sub-model describing systems of increasing complexity.

For example, the atom-atom potentials and the non-additivity corrections are obtained (*output*) from quantum-mechanical *ab-initio* computations, namely a sub-model, where information on nuclei and electrons are the only needed input. These potentials are subsequently used (as input) in a sub-model (specifically, statistical mechanics and Monte Carlo techniques) where we consider interactions between "particles" such as atoms, ions, molecules and macromolecules. At this level, *temperature* is introduced explicitly, thus also *entropy* can be introduced most naturally. Finally, the *time* parameter can be introduced—explicitly—and in such cases we are in the *molecular dynamics* modeling. Often we wish to start with a reasonable configuration of atoms and molecules and in this case we can use a Monte Carlo configuration with high probability as starting point for molecular dynamics simulations.

Parenthetically, we recall that the use of quantum chemical computations to obtain atom-atom potentials is today a viewpoint rather commonly accepted (for additional discussions, see Ref. 7 and Ref. 8 in this volume); this was not the case however, about ten years ago when we started to advocate the use of ab-initio computations[9] as an alternative to the—then predominant—use of empirical or semi-empirical potentials.[10,11] Equivalently our *"generalized model"* constitues a

point of view which is slowly becoming more and more accepted as the natural methodology needed to describe systems as complex as those found in biology.

Among the *outputs* from a Monte Carlo simulation we find *generalized potentials and forces* describing the interactions between two or more micro-systems,[6] each one composed of an ensemble of particles like atoms, ions, and molecules. Particularly interesting, are the "solvated ion"—"solvated-ion" potentials and the "solvated molecule"—"solvated molecule" potentials which can be used as input to represent the next level of complexity, as later shown in an application reported for the first time in this work, but proposed sometime ago.[6] As known, thermodynamics is in general assumed to be the sub-model which follows statistical mechanics, when we order the sub-models on a scale of increasing complexity. In our methodology, however, we can retain the use of statistical mechanics *and* Monte Carlo techniques for a *midy-system composed of an ensamble of both micro-systems and atoms, ions and molecules*. In our method, with the introduction of specific midy-systems, we retain the option to switch between the very detailed and specific statistical mechanical description of a micro-system and the very general but unspecific thermodynamical description of a macro-system. Note that *this option might allow to switch from a very short time scale*—typical of the molecular dynamics of the micro-system components,—*to a much longer time scale*. To our knowledge, such switching between time scales has not been previously considered in molecular dynamics (in general) nor in molecular biology (in particular). For reason of space we shall not expand further on our methodology; we recall that the "generalized model" here summarized is a natural follow-up of some very qualitative ideas concerning the organization of interactions of particles in increasingly complex systems, elsewhere outlined.[9-12]

Löwdin's generalized theory[13] could be taken as the theoretical foundation of our approach; on the other hand, since our postulates are simply input-output relationships relating sub-models, we need only an operational set of informatic-type postulates as foundation to our "generalized model". The chapter by Careri[14] is here noted, since it outlines scientific and philosophical bases leading to the need for some generalized model in order to describe biological systems.

To conclude this section we briefly recall few papers presented in this volume which are directly related to our work. In the paper by Palma[15] there is a very interesting link between our study on DNA and agarose; this points out the generality of some conclusions we have obtained in our specific computer experiments dealing with nucleic acids. In addition, in view of the above methodological discussion, it should not be necessary to point out the strong relationship existing between our work and the thermodynamical modeling, for example, by Anderson and coworkers[16] or in Manning's condensation theory.[16]

In the following, we shall define a *hydration site* and consider its steric, energetic, static and dynamic attributes. Then we shall extend the discussion to *counterions*

sites, and the closely connected subject dealing with iso-energy maps obtained from quantum-chemical potentials. Completing this, we are ready to discuss our recent Monte Carlo computer experiments.

Three-dimensional Models: Availability of Hydration Sites

Modeling the molecules of water hydrating DNA double helices has a rather long history, which starts with the wiring of physical three-dimensional models (either ball-and-stick or space-filling) attempting to obtain—by trials and errors—reasonable hydrating structures. We refer for example, to Lewin[17] and Forslind[18] as excellent examples of this methods of research, today superseded and mainly of historical interest. *Few observations should be made concerning these studies in relation to modern computer simulations.* As known, in the above physical models the atoms are generally represented by spheres and for each type of atom a pre-established number of connections to other atoms—representing bonds—are allowed. The "volume" of the atom is also pre-determined; Van der Waals radii, atomic and ionic radii and standardized bond-length have been used to represent "the volume of the atom" in the molecules.

Further, in wiring a molecular model, an atom is either *linked* or placed in *physical contact* or *separated* from another atom. Thus, it follows that in three-dimensional models, the assumed interaction energy, ΔE, between an atom A and a second atom B, is represented by square well potentials of type I—attractive—or II—repulsive—as given in Figure 1. Type I, corresponds to a chemical bond (the potential is modeled as an infinite repulsion in the region "a", followed by an infinite attraction in the region "b" and, thereafter, by a non-interacting region; type II approximates non-bonding interactions, in general, and intermolecular interactions in particular. From the form of the potential energy implied in the use of these "physical models", it follows that the information which can be derived when applied to hydrations, is essentially limited to predictions on the existence of hydration "site" and it is of binary type, namely, a "site" is either *available* or

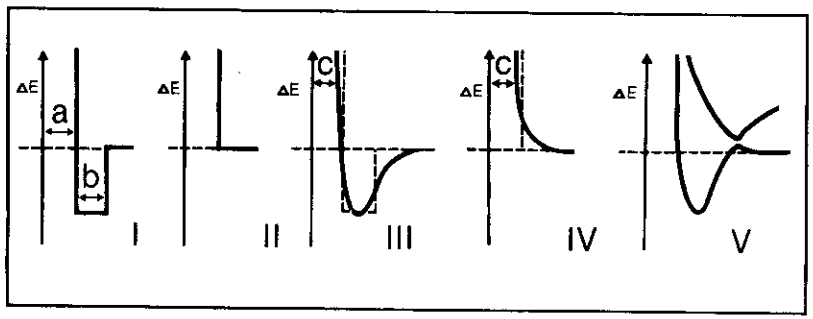


Figure 1. Different approximations to describe inter- and intra-molecular interactions.

non-available. The latter case occurs when an atom of the water molecule would have to "penetrate" within the volume (the infinitely repulsive region) of an atom of the solute.

Let us define a *hydration site*: it is a region of space with a volume and shape such as to closely contain a water molecule. A site position and orientation is generally defined relative to one (or more) nearby atoms of the solute molecule. "Nearby" means, in general, that the site is part of the first solvation layer of the solute. In an equivalent way one can define a hydration site with reference to atoms of the solvent. In these definitions it is tacitly assumed that a site is positioned and oriented in such a way as to be at an energy minimum. Clearly, one can not be sure to reach the energy minima when "physical models" are used.

Interaction Energy: Stability of a Hydration Site

As known the model potential I and II are only approximations to the model potentials III and IV of Figure 1; indeed an interaction potential is, generally, represented by a smooth, non-discontinuous function of the internuclear separations, either of attractive or repulsive type. We recall that more realistic representations of interaction potentials are given by the model potentials V of Figure 1, which, however, can be assumed to be merely a linear combination of interaction potentials of type III and IV. Note that we have over-imposed to the model potentials III and IV the corresponding square-well potentials in order to show the extent of the approximation in passing from models I and II to III and IV.

Since in models III and IV one assumes *finite* interactions, after a *hard core* region (designated as c), we are finally in a position to state, for any site, its *first* characterization, namely, *the site stability relative to the solute*: let us indicate it with $E_i(S)$, where i is an index to distinguish one site from any other site (capital S is used for the solute, whereas small s will be used for the solvent). From the sign of $E_i(S)$ we characterize a site either as *attractive* or as *repulsive* relative to the solute. Recalling that a solvent molecule is generally in contact with other solvent molecules, the next characterization of a site is its *stability relative to the solvent*, indicated as $\eta(s_i)$. This type of stability is obtained by summing all the interactions of the solvent molecule at site i, with all the remaining solvent molecules at the sites $1, 2, \ldots, i\text{-}1, i+1, \ldots, N$.

Let us recall that a water molecule can be approximated by a nearly spherical volume of about 30 Å^3. (A more accurate description depends on the specific type of "probe" selected to measure the water's volume; as shown later, a X-ray particle "sees" a volume which is not the same as the volume seen by a neutron or by an ion or a neutral molecule). In Figure 2, we report a cross section of the water-water interaction energy. This is accomplished by reporting an iso-energy map, obtained by placing a water molecule in the x-y plane (with the oxygen atom at the origin of the axes) and by computing its interaction with a second water molecule constrained to have its oxygen atom in the x-y plane (with $z=0$) and an optimal

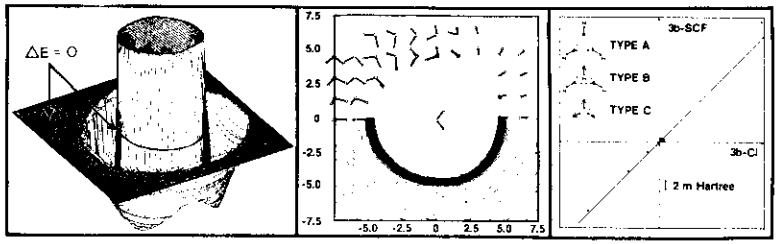


Figure 2. *Left*: water-water potential. *Center*: iso-energy map for water-water interaction energy. *Right*: three-body correction for water-water computed without (3b-SCF, Hartree-Fock limit) and with electronic correlation correction (3b-CI, configuration interaction).

orientation of the hydrogen atoms, such as to maximize the attraction between the two water molecules. In the left inset we report an iso-energy map in three-dimensions, in the central inset the same iso-energy map is given in two-dimensions. The inset on the right reports a comparison of three-body corrections obtained either with (three-body C.I.) or without (three-body S.C.F.) electronic correlation correction; this inset will be used later in a discussion on the water-water interactions in liquid water. Note that in the central inset, the water molecules have been shrunken by about a factor of two, to render the orientations more clearly.

To a first approximation the number, N, of sites in the first solvation layer is about equal to the number of spheres (with radius of about 3 Å) which one can pack all around the solute, forming a mono-molecular layer of water molecules. In Figure 3 the mono-molecular layer includes the water molecules 1 to 7; additional layers are shown. Even by inspection one can see, that, to a first order of approximation, the total interaction at the 4-th site is $E_4 = E_4(S) + E_4(s_3) + E_4(s_5)$ namely, $\eta(4) = \sim E_4(s_3) + E_4(s_5)$, (with $E_i(s_j)$ we have indicated the interaction energy of the water at the site i with the water molecule in the site j). By adding to η_4 the interactions with the waters 2, 10, 11 and 6 one obtains a very reasonable value for η_4. The interaction with more distant water molecules can be approximated with by some average interaction, volume dependent.

The statement that a water's site can be approximated to a sphere must not be over-extended, for example, by assuming that a water molecule can rotate around the sphere's center without an energy cost! As known, on each sphere we can identify regions with an *excess* and regions with a defect of electronic density, relative to the neutral atoms, yielding the dipole and the quadrupole moments, essential parameters for the understanding of hydration processes. The two minima, in the left and center insets of Figure 2, represent a hydrogen bond formation involving either a hydrogen atom or the oxygen lone pair of electrons, namely the positive and negative regions, respectively of the water molecule. In a somewhat old fashion chemical language, we would say that the rotation of a water molecule can bring about a breaking of those hydrogen-bonds existing between the water at

site i and the rest of the system. Notice that these "breakages" not only change the energy of the system but also its entropy.

Let us consider few limiting cases in the equation $E_i = E_i(S) + \eta(i)$ and let us define as $\eta(b)$ the average interaction energy of one water molecule in "bulk" water (at standard conditions); as known $\eta(b)$ is about 8 kcal/mol. For the case $E_i(S) = 0 \pm 1$ kcal/mol we have $E_i = \eta(i) \pm 1$ kcal/mol; if this condition holds for all the N sites of the first mono-layer, then the water molecules form (around the solute) a mono-layer with *clathrate type structure*. A very well known example of clathrate structure is the one of water molecules around methane,[19] shown in Figure 3, central inset. In these conditions for each molecule in the second layer and even more so for those in third, fourth, . . . , layers, $\eta(i)$ approaches $\eta(b)$.

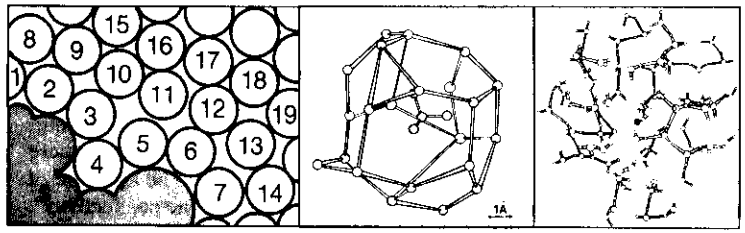


Figure 3. *Left*: Schematic representation of a solute molecule solvated by water molecules. *Center*: Clathrate structure of water molecules around methane from a Monte Carlo simulation at 300K temperature. *Right*: Water molecules and connectivity pathway for a K^+ ion from Monte Carlo simulations (300K).

For $E_i(S)$ to be *very* repulsive (more than 5 to 10 kcal/mol) it is most unlikely, under normal conditions, because of the Boltzmann probability distribution (this case would correspond to the "non-availability" of the site). This case therefore is not considered. The most interesting situations arise when $E_i(S)$ is attractive; we consider the sub-cases $E_i(S) \lesseqgtr \eta(b)$. For $E_i(S)$ approaching $\frac{1}{2}\,\eta(b)$ from zero, the site "i" will *likely* contribute to form a hydrating structure of clathrate type. For $E_i(S)$ more attractive than $\frac{1}{2}\,\eta(i)$, the water in the site "i" will likely prefer to select an orientation such as to optimize the interaction with the solute even at the cost of breaking hydrogen bonds with neighboring water molecules. An alternative statement is that the loss in $T\Delta S$ and of the water-water attractive interaction is more then compensated by the gain in internal energy. This is often the case for ions in water.[20,21,22] Obviously, when $E_i(S)$ is more and more attractive relative to $\eta(b)$, then the water in the site "i" will proportionally less and less experience the remaining solvent molecules. Note that $E_i(S)$ can be up to 10 times more attractive than $\eta(b)$, for example when one considers water molecules interacting with the phosphate groups in DNA or counterions. In Figure 3 (right side) we report the structure of water molecules around K^+: note that the K^+-H_2O interaction is about 3 times larger than the H_2O-H_2O interaction. In the insets, the hydrogen bonds are explicitly shown,

whenever two water molecules have the appropriate relative orientation and position. As below discussed, the connectivity pathways in the central and right insets are time dependent structures.

A question of basic importance in discussing solvent-solute interactions is related to the life-time of "the hydrating structures", like those reported in Fig. 3. If a structure has a lifetime as long as the mean lifetime for some event, like a reaction, then that structure is as real—for that event—as one determined for example, by X-ray diffraction studies, the standardly accepted time-scale for "chemical structures". In other words a "molecular structure" needs to exist only for some mean-time; the length of the "existence" of that structure is relative to some *"event-observation"* which can be as short as a chemical reaction.Thus the "X-ray" structures are like the top of an iceberg, showing only structures corresponding to very long life-times, but neglecting—at present—many more with shorter life-time but of basic relevance to chemical processes.

In literature a number of different terms have been used in an attempt to differentiate between hydration sites. For the case of a site yielding *clathrate*-like structures terms like *"hydrophobic"* and *"structure forming"* are often used. The term "hydrophobic", however, should not be related to the internal energy, E_i, of a site but, rather, to its free energy. For this reason the alternative term *"clathrate-site"* is here used, when we are mainly interested in the site's internal energy. The opposite of hydrophobic is *"hydrophilic"* but we shall use the alternative term "line-connectivity". In Figure 3, by comparing a "clathrate" and a "line-connectivity" site, we realize that a clathrate is a *"closed-surface connectivity site."* As elsewhere stressed,[20-22] ion-pairs in solution—like Li^+, F^-—exibit very pronounced connectivity pathways, originating in the strong field at the ion and stretching radially from it. The orientation of a water molecule relative to a field generating center, like an ion, is a function of the field strength at the water position; when an ion-pair is linked with a line-connectivity pattern, the orientations of the two water molecules at the two extrema (namely, at the two ions) are rotate by 180°. The tetrahedral coordination of water is particularly suited for allowing such rotations at low energy cost and with relatively short line-connectivity pathways. Note that the K^+ field is weaker than the Li^+ field and therefore, the line-connectivity aspect is less pronounced. In addition, notice that the representation of an ion as enclosed by more than one hydration shell, is equivalent to the line-connectivity representation, averaged over a sufficiently long time. Finally, notice the above given comments on the *life-time* of a hydrating structure. For additional discussions on hydrophobic interactions, see for example some of the references given in the chapter by Palma,[15] by Scheraga,[23] the note by Beveridge[8] and Ref. 6.

Local Displacements: Site's Probabilistic Description

As known, to a given temperature there corresponds a mean free path (and a mean rotational frequency) for the solvent molecules. This dynamical situation tells us that the previously reported representation of a site does not hold, if not as an

approximation. Only *locally stable* sites, namely those with $E(i) >> \eta(b)$, can be assumed to move about very little (in terms of space displacements) especially, if each one of the neighboring sites, j, has $E(j) \leq \eta(b)$ or at least $E(i) << E(j)$. The above condition characterize the site "i" as positionally very stable (deep energy potential and narrow enough as to fit only one water molecule). Note that sites which have deep potentials but are surrounded by other equally deep potential sites can move around considerably, namely, are locally un-stable. We assume that this is why in one of Dickerson experiments[24] the water molecules bound to the phosphate groups were at first not detected, even if these are at some of the *most attractive* sites in nucleic acids.[25]

From the above it follows that at a time t(2) a site can overlap with the position occupied by another site at time t(1). Let us count how *often* a site is occupied during a period of time Δt and let us do the same for all sites in the solution. We shall find out that for each position in space there is a *probability distribution* of finding a site. The *probabilistic* aspect for the site energy, position and orientation is its third characterization. As well known from statistical mechanics, one can obtain, theoretically, the probability distribution by following the prescriptions of a particularly powerful technique known as Metropoli's Monte Carlo method;[26] in two monographic studies[6,27] we have reported on liquid water and ions[27] and on biomolecules in solutions[6] and we refer the reader to these references for details.

It is well known that in ergodic systems we can obtain probabilistic distributions either by sampling the full phase-space (coordinates and momenta) or only the coordinate's space. In Metropoli's[26] techniques we consider only displacements. However, in *molecular dynamics* we solve the classical equation of motion for the particles of the system in a time-space framework. Molecular dynamics,[28] therefore, brings about the fourth characterization of an hydration site, namely its *mobility*. In summary, a site is characterized by its *availability, stability, probability* and *mobility*. The first two characterizations define the site's position, orientation and energy in a rather artificial model where temperature and often most of the solvent molecules are neglected; the last two characterization, probability and mobility, *restore realism* in the representation of hydration sites. The first two are typical outputs of quantum chemistry; the latter two are obtained by combining quantum chemistry with statistical mechanics. It is somewhat disconcerting to see how tightly some of today's "quantum biologists" hold to their "Schroedinger equation security blanket" without venturing into the post 1950's or 1960's developments; this viewpoint is corroborated by the realization that the progress in solving Schroedinger's equation for a truly large chemical system has been and is notably slow.

Sites for Counterions and Solvated ion Interactions

The basic macroscopic characterization of DNA at the electronic and electrostatic level is that DNA is a polyelectrolite. This fact, well appreciated since the early experiments on DNA, is the starting point for the thermodynamical studies on DNA over the last 10 to 15 years.[29] It is equally well known that the DNA fiber's diffraction

studies were at such low resolution that most of the sugars and of the bases, and all the DNA's hydrogen atoms and water molecules and counterions were not detected. Whereas the neglect of the counterions had a limited impact on DNA gross structural determinations, being X-ray data un-concerned with energetics, clearly an equivalent neglect is totally unacceptable in any theoretical models dealing with either the electronic states of DNA or the electrostatic potential or the electric field. Unfortunately, counterions *have been*, and even today *are* often neglected in quantum chemical studies dealing with the above properties. This is a rather serious situation; indeed the electronic characterization of DNA without counterions is as much acceptable as a characterization of the structure of an atom without its nuclear charge! As in an ionic crystal the position of the anions is dependent on the position of the cations, in an equivalent way, the position of the phosphates groups is dependent on the position of the counterions; the sugar group geometry controls the constraints on the possible conformations of the phosphates and—hence—indirectly of the counterions. However, let us now return to the problem of the *"sites"* for counterions.

Most of what stated in the previous section for an *hydration* site can be restated for a site occupied by a counterion. In DNA, all the hydration sites which are *strongly* stable and with an orientation such that the water's hydrogen atoms point towards DNA, whereas the oxygen atom points away, are good candidates to be also counterion sites. This is obvious on the bases of rather elementary electrostatic considerations. The range of the ion-ion interaction deserves some considerations. Whereas, the water-water interaction is nearly zero at an oxygen-oxygen distance of 9-10 Å, two counterions will strongly repel each other at these distances. Ion's hydration, however, cuts down the repulsive interaction; indeed the water molecules solvating an ion not only add stabilization to the system (the ion-water interaction) but also screen the ion-ion repulsion. This screening is very important for the understanding of the role of water in DNA and can be nicely simulated with Monte Carlo computer experiments. In the following, we consider the repulsion of two cations in water in some detail.

The *"solvated ion"-"solvated ion"* interactions can be computed by considering many water molecules and two ions, for example two K^+, placed inside a sphere of radius R; the two ions are positioned, symmetrically displaced, from the center of the sphere by an amount L, where $L \ll R$. In our experiment R is 15Å and $L \leq 5$Å, namely the ions are always far away from the boundaries of the sphere (note that our selection of a sphere as a volume, allows for easy corrections of the bulk water's field outside the sphere). For very large separation of the two ions, we switch to a different Monte Carlo realization, where we consider two spheres of radius $R' > R$, each one with an ion at its center; R' is such as to contain a number of molecules of water larger than the computed coordination number.[21,22] The computed interaction energies, at short and long ion-ion distances, are then used to obtain the fitting constants for an analytical expression of the solvated ion—solvated ion interaction; the expression has the same analytical form as the one selected to represent the bare ions interaction. We recall that the latter is obtained from near to Hartree-Fock energies computations on two bare ions.

As expected, the *solvated ions repulsions are weaker than the bare ions repulsions*; however, the rather popular notion that the decrease can be equated to a constant factor (related to the dielectric constant) is obviously incorrect, as seen by inspection of Figure 3, inset at right side. The line-connectivity pathways, namely, the filaments (or the threads) of hydrogen bonded water molecules, generated by the central field of one ion are expected to be severely perturbed by the presence of a second ion. The *distance-dependent perturbation* brings about a strong an-isotropy in the *line-connectivity*. Note that the connectivity pathways play an important role in differentiating *static* from *dynamic* situations. Indeed the connectivity pathways have average lengths, directions and lifetimes, which are function of the static (time-independent) field strength and of its dynamic (time dependent) fluctuations and/or modulations. From the above observations one can express some perplexity for those models which rely on "cavities" carved into the solvent and around a solute, starting immediately at the solute border. *Most clearly the cavity should be large enough as to include both the solute and its connectivity pathways*. Note, in addition, that every solute molecule, polar or apolar, is enclosed by *non-bulk water at least in its first shell*. The interaction energy between two "solvated molecules" is expected to be expressed by potentials of standard form, for example as given in Figure 1. Indeed, as the interaction potential between two atoms is the result of the electron's mobility around the nucleus, perturbed by the second atom, in an analogous way the interaction potential between two "solvated molecules" is the result of the water mobility around the solute, perturbed by the second "solvated molecule." Note, however, that the time-scales for the two interaction potentials are vastly different. In this view-point, the double-well potential found for the interaction energy of two methane molecules in a water solution[30] and later re-analyzed in a note published in this volume[8] represents a somewhat "expected" result in agreement, for example, with the physical model of Figure 21 of Ref. 6.

The decrease in the electrostatic repulsion due to the water molecules located in between two cations, is a general phenomenon for ionic-type charges separated by a dielectric and—therefore—equivalent findings are expected for a pair of phosphate groups—negatively charged.

In conclusion, from the above general discussions few of the fundamental aspects in the hydration of DNA have become apparent: by adding water to DNA the phosphate-phosphate repulsion is decreased by screening, therefore, stabilizing the DNA system; at the same time counterions condense around DNA, further decreasing the phosphate-phosphate repulsion and bringing about electrostatic stablization. *Secondly*, water molecules solvate the counterions, thus decrease the ion-ion repulsion, which would have destabilized the system. *Further*, since ionic interactions are very long range, one expects to encounter in the system "DNA + water + counterions" collective effects, very prominently. *Finally*, since the interaction energy of cations to water is very specific (for each cation), we expect that specific and local effects will complement the collective effects. As shown elsewhere,[6] these effects are essential factors in explaining, for example, the ion-selective transition from one DNA conformation to another.

Iso-Energy Maps for DNA's Components

The iso-energy maps are very instructive for an *initial* understanding of the interaction between water and DNA (or its components) or between ions and DNA. Let us assume that we are in position to compute, reliably and very rapidly the interaction energy between a molecule of *solvent* and one of *solute* (we shall later turn on this point in detail, when we shall comment on our interaction potentials). Then, we can generate readily iso-energy maps; below we recall the technique to obtain iso-energy maps between a solute and a molecule of water. Let us consider a plane relative to the solute (for example, a plane bysecting the solute molecule, or if the solute is a planar molecule, the molecular plane itself) and let us over-impose on the plane a grid with a square mesh. At each grid point we perform the following operations: 1) we place the nucleus of a water's oxygen atom, 2) compute the water-solvent interaction energy for as many orientations of the water's hydrogen atoms as necessary to determine the most stable interaction energy. After this process is repeated at each grid point, we connect the positions of equal value for the interaction energy with a contour-line, called the iso-energy contourline. An iso-energy map is obtained by considering all the iso-energy contours, with a predetermined contour to contour energy difference; for additional comments see references[6] and[31]. An iso-energy contour map for a water molecule interacting with a second water molecule has been previously shown in Figure 2, in the inset at the center; the optimal orientation assumed by the second water molecule at different sites is shown explicitly in the upper part of the symmetric map. If we consider an iso-energy map not for a water molecule, but for an ion, the computational process is essentially the same; since, however, no orientational optimization is needed, the computation is notably faster.

In the last several years this graphical technique has been also adapted to display *electrostatic potential maps*, following the proposal of Scrocco and Tomasi.[32] The electrostatic potentials are by now very popular because of the easiness to obtain the corresponding maps. It should be stressed, however, that the *electrostatic* potential maps, since *by construction* limited to represent the electrostatic interaction of a *point charge* with the electronic density of a molecule, *neglect many effects, like polarization, charge transfer, dispersion and exchange contributions to the total interaction energy*. As a consequence, for example from the electrostatic approximation one cannot obtain specificity for the interactions of Li^+, Na^+, K^+ with DNA; further the electrostatic approximation *fails* when the point charge marginally overlaps with the electronic density distribution of the solute and *fully breaks* down when the point charge substantially overlaps the density distribution, namely, in the example of the above cations, near equilibrium and at shorter distances. *Electric field maps* have been (and are) used[33] to complement electrostatic maps; note, however, that the field maps have the same limitations as the electrostatic maps.

In Figure 4 we report the iso-energ maps for two separated bases guanine, G, and citosine, C, and for the base-pair, G-C, placed at the left, center and right inset,

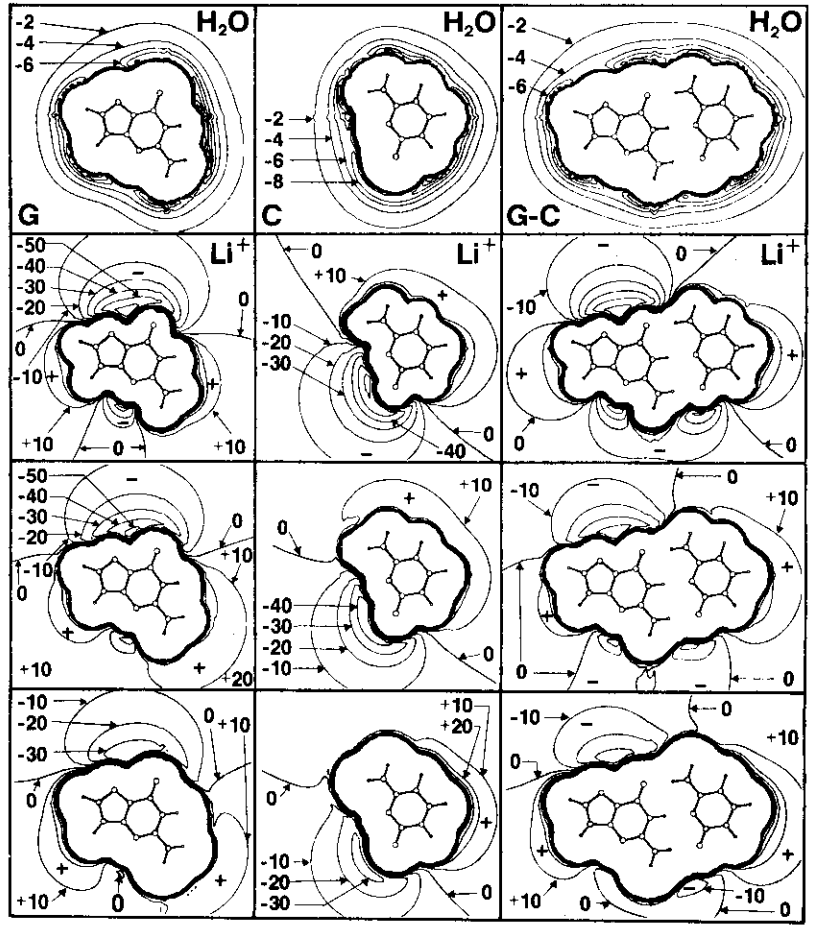


Figure 4. Iso-energy maps for G (left), C (center) and G-C pair (right) interacting with water (top line), Li^+ (second line), Na^+ (third line) and K^+ (bottom, or fourth line of insets).

respectively. In all cases the iso-energy map is for the plane containing the molecular plane. Starting from the top, the first line of insets reports iso-energy maps with one molecule of water; the second, third and fourth (bottom) lines reports maps of interactions with a Li^+, Na^+ and K^+ ion, respectively. The iso-energy maps for G, C and G-C given in Figure 4—top line—are graphically superior to those previously reported.[34] In these maps the contour to contour energy difference is 2.0 kcal/mol for the case with water, and 10.0 kcal/mol for the case with the ions; repulsive

contours above a given threshold have been omitted. One can easily see the attractive regions around the hard core; in order to help visualization, we have explicitly indicated the attractive (negative sign) and the repulsive regions (positive sign).

The strongly repulsive region, which starts at the zero kcal/mol contour and rapidly increases in value, when one moves toward the inside of the molecule, marks a shape which delimits and *defines* the hard core. As one can verify by inspection, the shape of the hard core is different when obtained from the interactions with a water molecule rather than with a counterion. Would we have computed the C, G and C-G iso-energy maps not for water but for some other neutral molecule, again the shape would have been somewhat different; and the same conclusion would be obtained if we would have considered the interaction of our bases and/or base-pair with either X-rays or with neutron-beams. Sometimes we tend to forget that the *structure of a molecule is different for each different probe*. From these maps it is also very obvious that a substituted G and/or C (like a methyl or ethyl group substitution) would bring about very different maps.

The attractive and repulsive regions are *qualitatively* different in comparig the maps for water with those for the three ions; in addition, *quantitative* differences are present in comparing any one of the three ions with another ion. Because of the small size of the insets, we note that in addition to the easily visible differences, the size of the molecular core of G, C and G-C decreases from K^+ to Na^+ to Li^+ by an amount proportional to the corresponding ionic radius.

One more point is of notable interest: *the base-pair G-C attraction to the ions occurs at positions corresponding to the major and minor grooves* in DNA, whereas the repulsion to ions occurs at two sides of the base-pair in correspondence to positions occupied by the sugar units. Thus, the base-pair seem particularly tailor-made to interact with ions when connected to the sugar-phosphate groups. But this is only one part of the story! As reported elsewhere[35a] the sugar unit in DNA *donates about one quarter of an electron to its base*: thus each base is negatively charged and the G-C base-pair has, in total, about one half of an electron in excess, acquired by the transfer. Obviously, because of elementary electrostatic considerations, *this charge transfer notably increases the attraction of a counterion to the bases and base-pairs and at the same time makes the base-base interaction somewhat less attractive, namely* weaker, *because of the charge transfer induced electrostatic repulsion*. The ion-base increased attraction (due to charge transfer) can be seen by inspection of Figure 5, where we compare the iso-energy maps for G interacting with Li^+, Na^+ and K^+ in three different conditions, namely G either as a neutral molecule (separated from DNA) or as part of B-DNA or of Z-DNA.[35a] It is noted that the computation of the charge transfer in B-DNA and Z-DNA were obtained by considering an olygomer with three sugar and two phosphate units and two bases, G and C, either in a conformation corresponding to B- or to Z-DNA.[35a] It is further noted that these computations correspond to the most extensive ab-inito computations ever performed in quantum-chemistry and quantum-biology; for details we refer elsewhere.[35a] Finally we recall that the charge transfer has also been

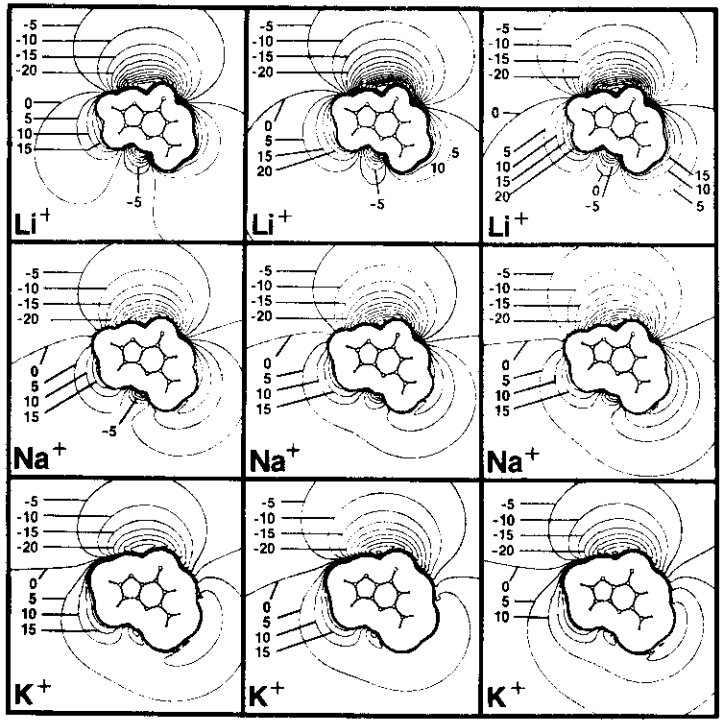


Figure 5. Iso-energy maps for G with Li^+, Na^+ and K^+. The base has charges either as in separated molecule (left), or as when in B-DNA (center) or as when in Z-DNA (right).

obtained in ab-initio computation of the electronic bands in B-DNA[35b,36] using crystal orbital techniques,[36] however in the latter computations[35b,36] due to numerical complexity, one had to introduce simplifications not present in the molecular computation on the B-DNA and Z-DNA fragments.[35a] This point is discussed in Ladik's work reported in this volume.[37] Parenthetically, the above overall electrostatic model is notably different from the one, where both the counterions at the bases, and the charge transfer from the sugar units are neglected;[38,39,40] this difference in modeling becomes more and more wide after our inclusion of solvent effect, non-zero temperature and statistical averaging.

In previous papers[41,42,43] we pointed out that the counterion base-pair strong attraction is not limited to a counterion positioned to the base-pair molecular plane, but is present also when the counterion is above the base-pair molecular plane. In this geometrical configuration the counter-ion can attempt to penetrate the base-pair in the hydrogen bonds region and influence the base to base relative separations and

orientations. This is tantamount to state that the counterion can influence base-pair opening, the relative angular orientation of the two bases, their relative distances, namely all the vital elements of nucleic acid conformations. In Figure 6 we consider three-dimensional iso-energy maps for Li^+ and K^+ interacting with G-C either in the G-C molecular plane (z=0. a.u. first and second line of insets) or above it by 3.0 a.u. (z=3. third and bottom lines of insets); in addition, the two bases are either at

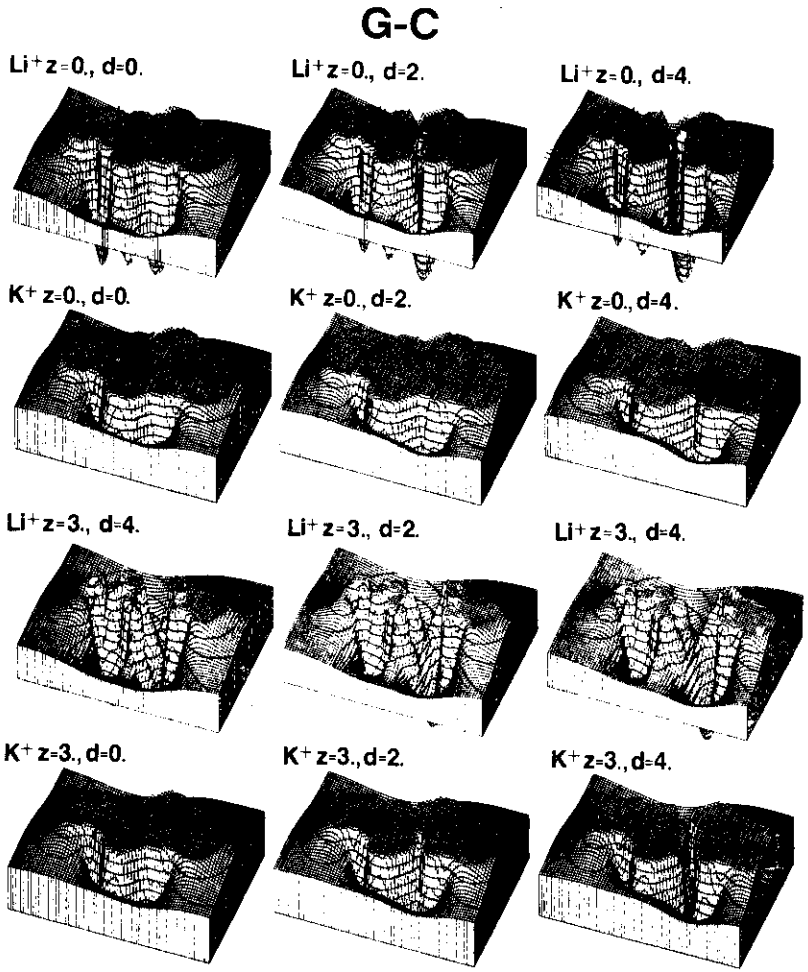


Figure 6. Three-dimensional maps of G-C with Li^+ and K^+ at z = 0 a.u. (top and third rows), z = 3 a.u. (second from top and bottom rows) for d = 0 a.u. (left column), d = 2 a.u. (center) and d = 4 a.u. (right).

the standard (equilibrium) base-pair geometry (d=0.0 a.u.) or are analyzed after increasing the distance by either 2.0 a.u. or 4.0 a.u. (d=2.0, d=4.0, respectively). The two-dimensional maps for G-C in Fig. 4, specifically the right side insets on the second and fourth (bottom) lines, corresponds to the two insets of Fig. 6 designated as Li^+, Z=0., d=0, and K^+, z=0., d=0.; in Fig. 6 the horizontal lines are iso-energy contours. We can see (look at the big canyon in between the two bases, top line, right insert) that Li^+ can penetrate in between the two bases even when it is constrained to stay in the molecular plane as long as the bases are pushed apart a bit (z=0.0 a.u. and d=4.0 a.u.). For K^+ this does not occur, because K^+ is less attractive to the bases. If we constrain Li^+ and K^+ in a plane 3.0 a.u. above the base-pair, then both ions are attractive when immediately above the hydrogen-bonding region when d=4.0 a.u. For Li^+ this attraction remains even at d=2.0 a.u. and d=0.0 a.u. Notice, that the field of the phosphates will favor the above processes, as one can easily surmise by inspecting the iso-energy reported for example, in Figure 70 of Ref. 6.

Thus, we have shown that the counterion can penetrate very near or be in between the two bases of a base-pair, bringing about an attractive energy which balances the loss of energy due to breakage of the hydrogen-bonds. To us it is therefore difficult to reconcile this finding with the rather popular neglect of considerations about the counterions effects in discussing intercalations, defects, kinks, mutations, base-pair stability, etc., namely, the vital processes of DNA reactivity and dynamics. Our prediction on the position of counterions very near the bases,[41,42,43] the above findings concerning the intrinsic attraction for counterions by the base-pairs at those positions corresponding (in DNA) to the grooves, the additivity of the phosphate's fields which reaches a maximum at the region near the central axis of DNA,[6,43] and the known ability of counterions to bring about conformational variations point clearly to the need of a model for DNA where the counterion position is as important as the phosphate position.

What we have stated above for the G and C bases and the G-C pair could now be restated for the A, T and the A-T base-pair, as clearly shown by the insets of Figure 7. Notice, however, the quantitative differences for the G-C and A-T base-pairs. The interactions of A, T and A-T with water are not reported, since elsewhere available.[34]

We conclude this section by recalling Sandaralingam's long standing identification of the central role played by the sugar unit in the determination of the structural conformation for DNA:[44] those experimentally based deductions, are here complemented by our theoretical modeling, where the charge transfer from the sugar to the base is seen as a primary process, notably enhancing the attraction of counterions to the bases and/or base-pairs and—at the same time—weakening of the base-pair interaction. Thus the sugar unit critical role for controlling conformational stabilization emerges also at the electronic structure's level. There are two main locations for the constraints of the double helix phosphate groups dynamics, one at the sugar units, and the second one at the base pairs: *the charge transfer weakens*

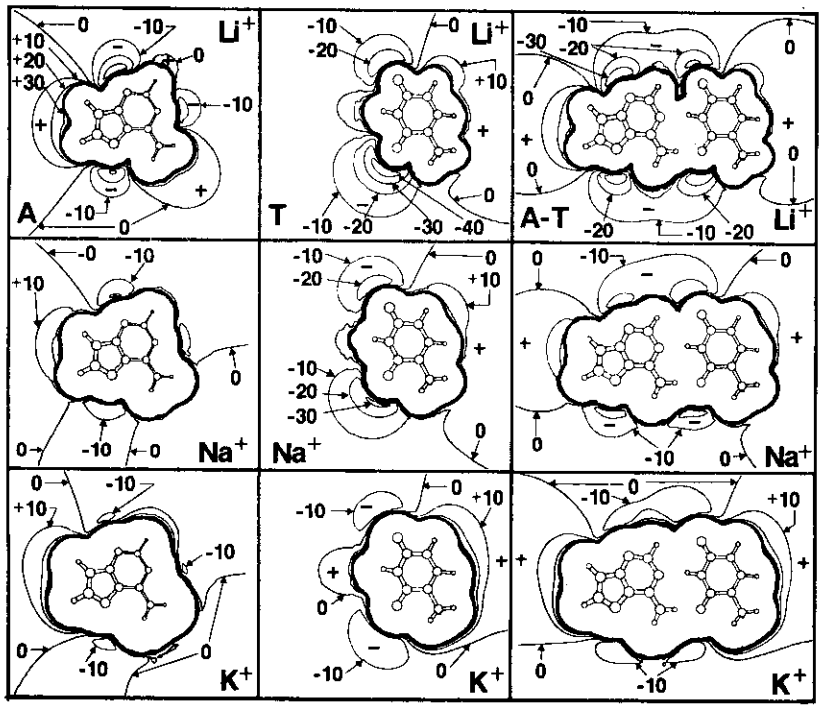


Figure 7. Iso-energy maps for A (left), T (center) and A-T base-pair (right) interacting with Li^+ (top), Na^+ (middle) and K^+ (bottom).

the base-base attraction, thus it weakens the constraints at base-pair, and adds flexibility to the polymer.

Iso-Energy Maps Of Water With B- and Z-DNA

A few years ago we reported preliminary Monte Carlo simulations of a few water molecules interacting with A-, and B-DNA and the corresponding iso-energy maps.[45] In those early studies, the atom-atom pair potentials did not include the charge transfer from the sugar to the base units. Below we shall analyze a few iso-energy maps for a water interacting with B- and Z-DNA; in these maps the charge transfer is included in the interaction potentials. The maps of Figures 8 and 9 are obtained by flattening out the iso-energy map corresponding to a surface of a cylinder of radius R, co-axial to DNA long axis. The maps are computed for a fragment of DNA composed of three-full DNA turns (for example, containing 30 base-pairs in the case of B-DNA); only the central turn is reported in the figures. To help visualization, each map is reported twice, once in two-dimensions (right) and once

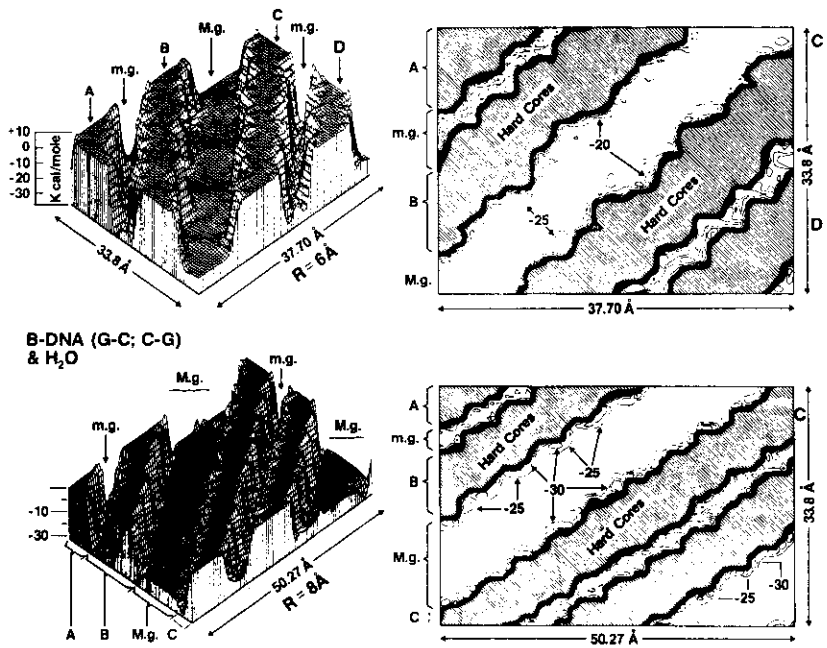


Figure 8. Cylindrical iso-energy maps for B-DNA with H_2O.

in three-dimensions (left). In the two-dimensional representation, the ordinate corresponds to the length of one full turn (namely 33.8Å for B-DNA and 44.58Å for Z-DNA) and the abscissae corresponds to 2π R, where R is the value selected for the cylindrical map's circumference. In Table I, we report for a few R-values, the DNA atoms bisected by the cylindrical surface (or very near to it) for the case of B-DNA and Z-DNA: notice that these atoms constitute the hard core shown in the maps. As done in Figs. 2 and 6, the core repulsion energies are reported up to a given value: this brings about the "mesa" like appearance of the hard core regions. Clearly, in B-DNA the smaller the value of R, the larger the hard core region and correspondingly the smaller the grooves region (major and minor grooves are designated as M.g. and m.g., respectively). This is not true for Z-DNA, as clear from the data reported in Table I. The designation, B-DNA (G-C; C-G) is introduced for a B-DNA sample with a base-pair sequence G-C and C-G; the designation B-DNA (G-C; A-T) is for a B-DNA sample with 50% G-C and 50% A-T, elsewhere reported.[42,46] In previous papers only the B-DNA (G-C; A-T) sample was considered, thus the above differentiation in the designation was unnecessary.

For B-DNA (Figure 8) the M.g. valley's floor is at about -15 kcal/mol for R = 6Å, the values of -20 and -25 kcal/mol being found only near the edges of the valley,

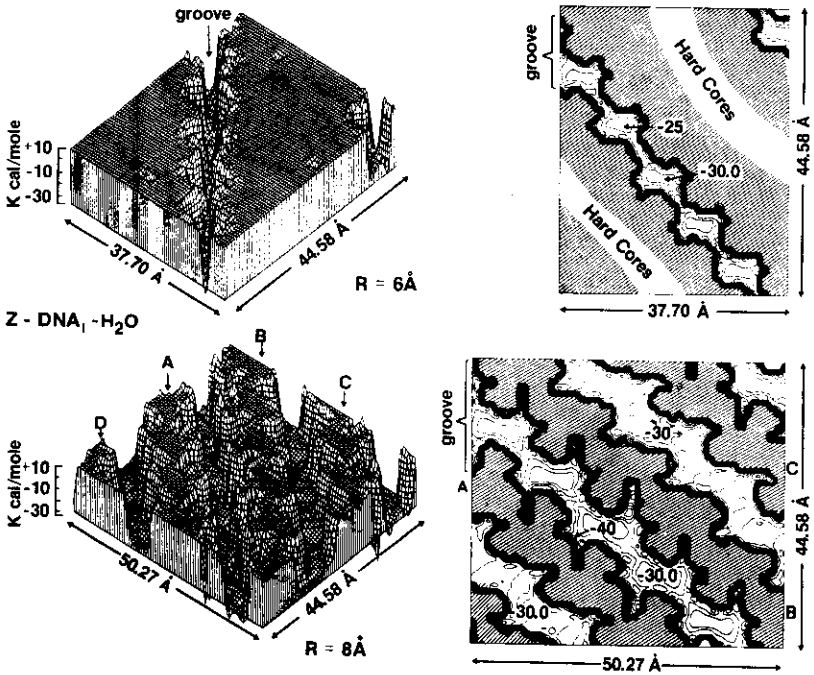


Figure 9. Cylindrical iso-energy maps for Z-DNA with H_2O.

immediately before the beginning of the hard core ridges. The m.g. floor is energetically more attractive to water and—as expected—much more narrow. These findings available for some time[45] have been *lately* confirmed by an experimental verification.[3] For R = 8Å we note the expected—in B-DNA-decrease in the width of

Table I
Atoms Participating to the Hard-Core in the Cylindrical Maps.

	B-DNA (G-C,C-G)
R = 6	: O1′, C1′
R = 8	: O3′, P, O2P, O5′, C5′, C4′, C3′, C2′
R = 10	: O1P
R = 11	: O1P
	Z-DNA
R = 4	: O1′C, C3′C, C2′C, C4C, C5C, C6C, O3′G, N3G, C2G, N1G, C6G
R = 6	: O1PC, O5′C, C5′C, C4′C, N4C, PG, O1PG, O2PG, C5′G, C3′G, C4G, O6G, C5G
R = 8	: O3′C, PC, O2PC, O5′C, C4′G, O1′G, C2′G, C1′G, N9G, C8G, N7G
R = 9	: O2PG

the hard core "mesa". (Note that the hard core has been truncated to +10 kcal/mol providing the "mesa" appearance). For more details we refer to the many maps previously reported.[6,45,46] For Z-DNA, the maps are very different, as shown in Fig. 9. We recall that the one-groove map at R = 6Å evolves, at R = 8Å, in an "apparent" two-grooves map with one groove deep and a second more shallow; the latter corresponds to those atoms which are on the exterior of Z-DNA. Only the deepest groove survives at smaller R values. Once these figures are reported for B and Z-DNA, *it seems somewhat redundant to point out that such large differences in the counterion DNA interactions, are bound to bring about notably different hydration patterns in B-DNA relative to Z-DNA and, consequently, energetically different stablizations*[46] *in the two conformers.*

Iso-Energy Maps for Li^+, Na^+, K^+ With B- and Z-DNA

In Figure 10 we display the iso-energy maps for K^+ with B-DNA (G-C; C-G) at R = 6, 8, 10 and 11Å. Starting at the larger R value (R = 11Å) one sees the small core regions corresponding to the oxygen atoms O1P at the phosphates (see Table I), then a larger cross section of the individual phosphate groups (R = 10Å), and then the large valley of m.g. and the larger M.g. valley (R = 8, 6Å). Comparing the maps for R=6A and R=8A with the corresonding one for the water's iso-energy (Figure 8), one realizes that *K^+ can have higher mobility than H_2O.* The energetic values reported at the valley floors in Fig. 10 are extremely large, over thousand kcal/mol: rather clearly these energies are totally unattainable in a laboratory experiment on DNA since of the order of plasma-type energies. Indeed in no experimental condition three full turns of B-DNA could be stable, when only one counterion (or one charge) is present. For this reason the iso-energy maps offer a useful but preliminary and qualitative picture; the same holds for the electrostatic map or for field maps, when obtained in equivalently "unrealistic" conditions;[38,39,40] we recall that (in our opinion) in such conditions, quantitative comparison of *small energetic details* (5 to 10 kcal/mol relative to a total of about 1000 kcal/mol) and differences between DNA conformers should not be taken too seriously!

A very interesting feature can be noted at R = 6Å: the K^+ counterion can penetrate at the base and is "nearly" able to pass across grooves, for example from the major to the minor grooves. This observation (this time obtained for a DNA fragment) will turn out to be very important, because it shows that a counterion can come very close to the hydrogen bonds of the base-pairs and therefore interfere with the base-to-base bonding as previously discussed in reporting Fig. 6; note that we have previously called attention on this mechanism.[42,46] Note in addition the non-large differences in the energy at the valley floor from R = 6Å to R = 11Å; this is the main reason for the extended mobility in the x-y plane (perpendicular to the main DNA axis) for the Na^+ counterions.

In Figure 11, we report the iso-energy maps for K^+ interacting with Z-DNA at R = 4, 6, 8 and 9Å. The existence of only one groove is evident at R = 4 and 6Å; at R = 8 and 9Å we approach the external atoms of Z-DNA and an "apparent" second

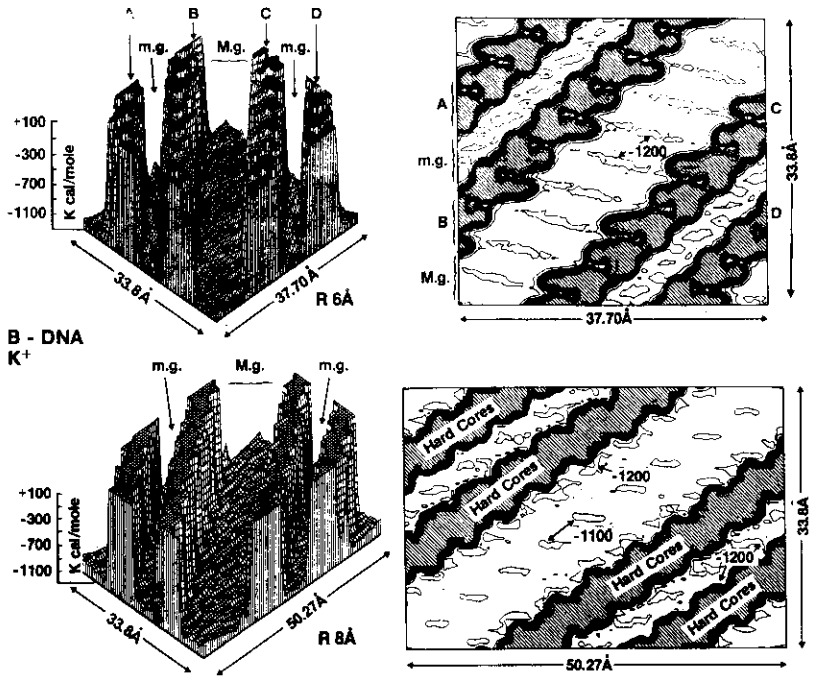


Figure 10. Cylindrical iso-energy maps for B-DNA with K^+ at R = 6, 8, 10 and 11Å.

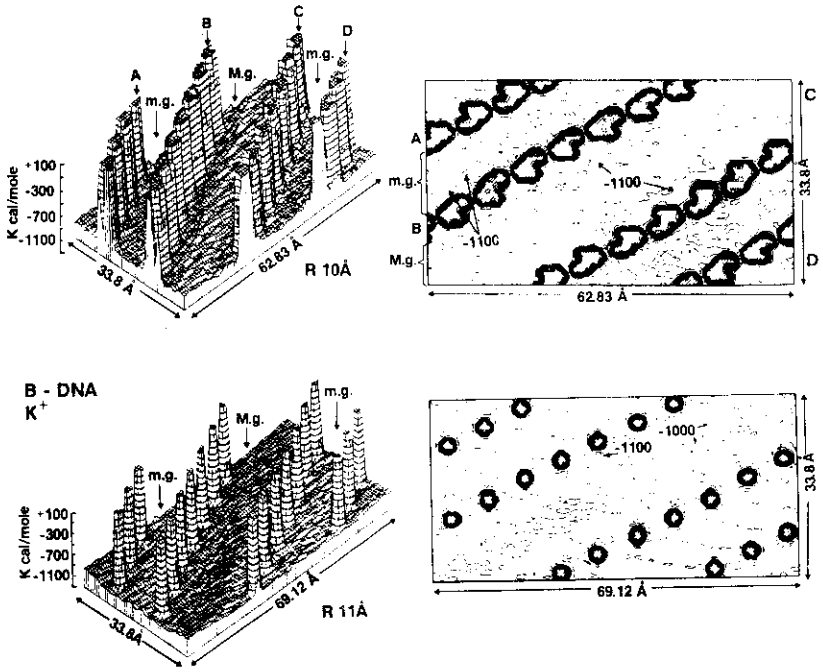


Figure 10 continued from previous page.

groove is shown, responsible for the "many cusps" shown in the hard core regions at R = 6 and 4Å. The possibility of groove to groove (x-y plane) migration can not be ruled out for R values larger than about 7Å; this feature is present also in B-DNA. Finally, in Figure 12 we compare for B-DNA (G-C; C-G) the iso-energy maps for Li^+, Na^+ and K^+ at R = 6Å. The valley floor energy is smaller (more attractive) in the order Li^+, Na^+, K^+. A more detailed map reported in an IBM-Research Report, which complements this work,[47] shows these energetic features in detail. By inspection of Figure 12, we would like to conclude that among the notable differences between the counterions is the ability to penetrate from one groove to the next one.

Intermolecular Potentials

The intermolecular potentials used in this work are those describing the interactions between a molecule of water and a second one or a counterion or DNA and those for a counterion with a second one or DNA. The interaction of either a water molecule or a counterion with DNA is obtained from ab-initio computations of the interaction energy of a water molecule (or a counterion) with the *components of DNA*, namely the bases,[48] the sugar units[49] and the phosphate units.[25] To include the charge transfers in our potentials, *we have used the atomic net charges obtained from the large ab-initio computation*[35a] *previously discussed*. Notice that in our interaction potentials, the atomic net charges are explicitly present in the analytical expression representing the atom-atom potentials; therefore, the substitutions of the value of the atomic net charges from computations on the separated components of DNA (thus with no charge transfer) with those values obtained from the computation with charge transfer is an easy task.

Concerning the reliability of our potentials we recall few verifications performed over the last several years, some summarized in Figure 13. *Firstly*, however, we stress that it is practically impossible to obtain potentials very accurate and *equally reliable for every possible observable*. Indeed a potential will always reproduce some experiment more reliably than some other. *Secondly*, we recall that our potentials have been constructed with the aim to reproduce interaction energies and structural data and to be used mainly in the contest of Monte Carlo simulations. Here, we should point out that our water-water potentials have been used quite

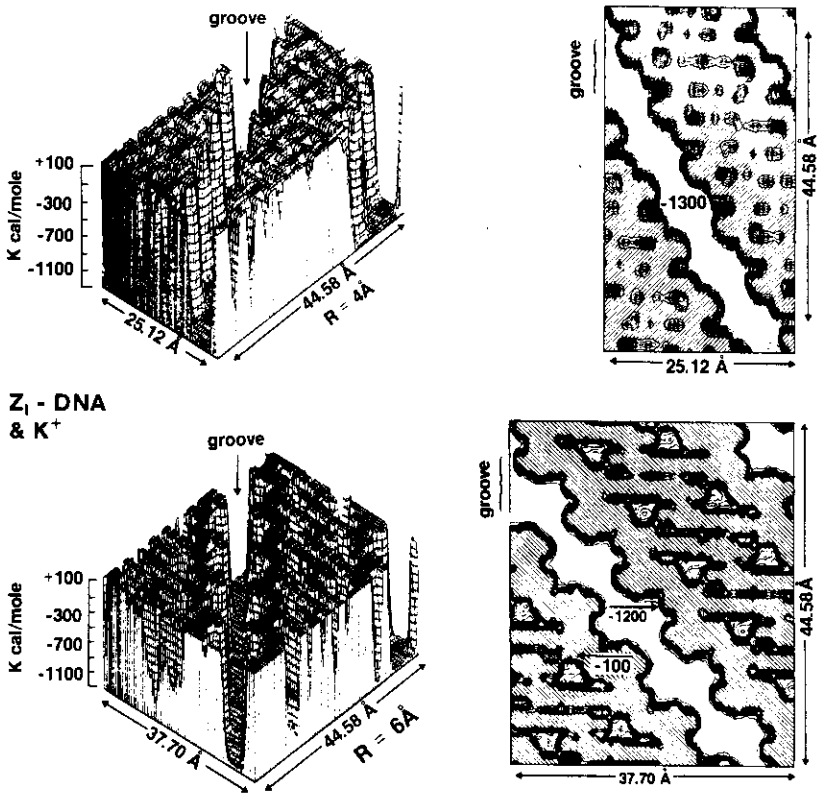


Figure 11. Cylindrical iso-energy maps for Z-DNA with K^+ at R = 4, 6, 8 and 9Å.

successfully also in molecular dynamics computations by Scheraga and his collaborators.[50] *Thirdly*, we recall that the atom-atom potentials obtained by considering two-body interactions (that is, only two molecules) can be defective, *since the three-body interactions* (which is not included) *can be relatively large*. We have stressed this point for a number of years and we refer to Ref. 6 for a general discussion. Specific applications have been considered, in particular, the three-body corrections for three molecules of water,[43,51] for two molecules of water and one ion[52] and for one molecule of water and two ions.[53] A new and very extensive work on the non-additivity correction is now nearly completed, *yielding for the first time a notably accurate two-body potential*, and, *a sufficiently accurate three-body potential for the water molecules*; the former is obtained at about 0.2 kcal/mol resolution, the latter is obtained at about 0.4 kcal/mol resolution (few hundreds ab-initio computation with extended basis set and very large configuration interac-

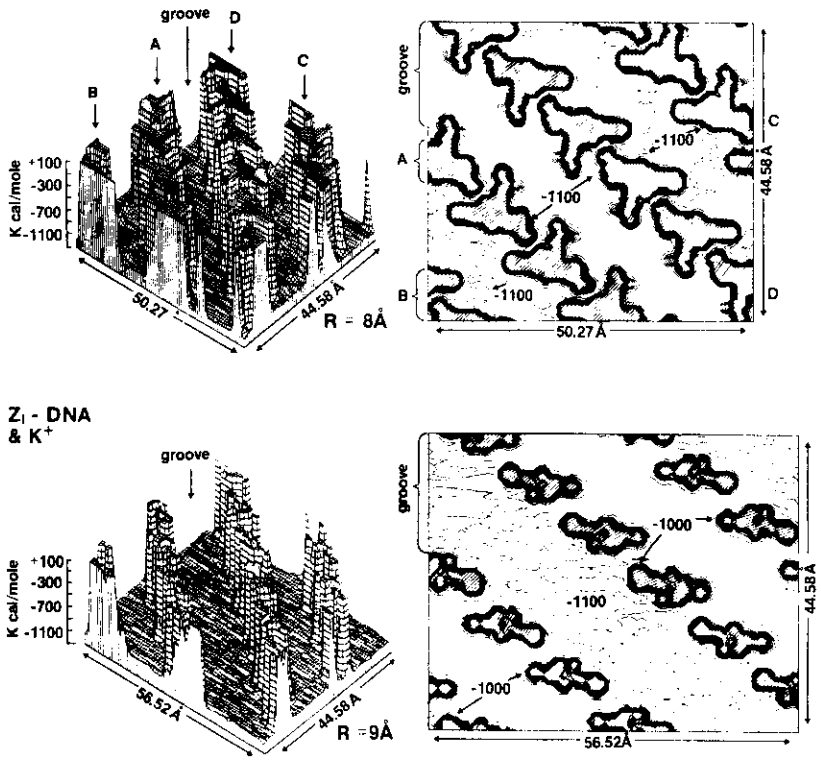


Figure 11 continued from previous page.

tion were needed). *Lastly*, we recall that in dealing with large systems, like those we have been working on for the last decade, one has to balance many opposing factors and attempt to account for first order corrections, neglecting smaller ones; thus our overall strategy is not the same one we would select would we be interested, only in a much simpler system, for example liquid water. Figure 13, where we compare simulated results with few experiments, indicates we have reached a reasonable balance; for more details see Reference 6.

On the two insets to the left we compare the X-ray and the neutron-beam scattering intensity for liquid water computed either with our two-body water-water potentials or obtained from laboratory experiments; the temperature is 300 K. In our simulation we have used the experimental value of the density of liquid water. Would we have used a notably smaller value for the density of the liquid, we would have

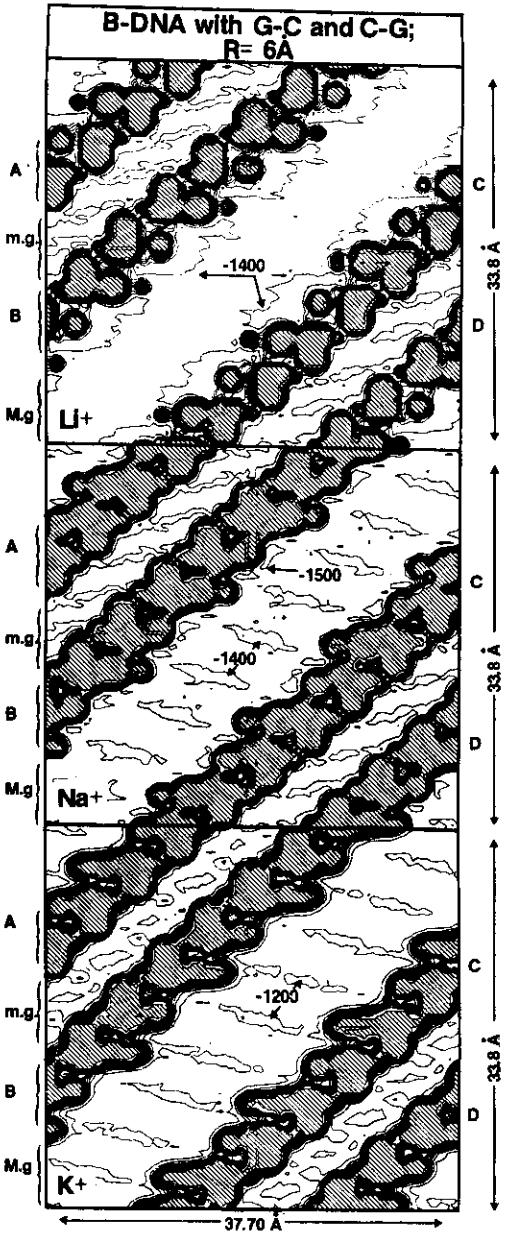


Figure 12. Cylindrical iso-energy maps for B-DNA and Li^+, Na^+ and K^+ at R = 6Å.

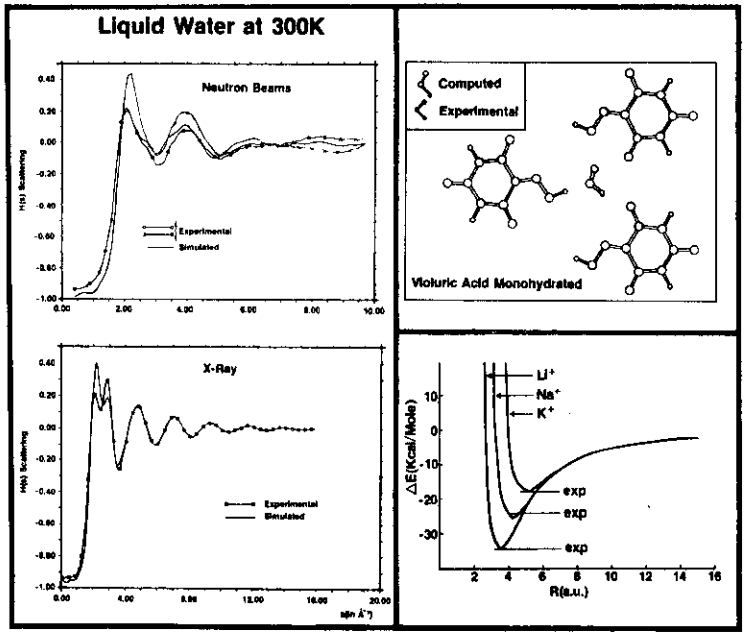


Figure 13. Comparison of simulated and experimental data.

certainly obtained different scattering intensities. Note that the agreement obtained with this experiment is likely the best one reported thus far in literature.[27] Note also, that the neglect of three-body correction represents the main source of errors in our simulation. On the two insets on the right we report the position and orientation of a water molecule in a monocrystal of mono-hydrated violuric acid. The latter is of interest because of the availability of high resolution X-ray and neutron-beam diffraction.[54,55] The experimental position and orientation essentially coincide with those from the theoretical computation (for additional details including comparison of Debye-Wolles factors, we refer elsewhere[56]). Finally, in the bottom inset to the right, we compare computed with experimental energies for a water molecule interacting with Li^+, Na^+ and K^+. Again, the notable agreement is evident. From these verifications and from many additional ones not reported here but available elsewhere,[6] we conclude that our potentials are certainly sufficiently reliable for the task for which they were constructed and used, as reported below. One more note: our potentials have been constructed to be transferable[27,48] from molecule to molecule: we note that in the last few years the notion of constructing "transferable potentials" is becoming more and more popular.

We conclude this section by stating that a) continuous refinements are needed, particularly the inclusion of many-body terms leading to a fuller account of polari-

zation effects, b) alternative but realistic ways to construct potentials are welcomed and, c) the use of ab-initio computations as the starting reference data for obtaining intermolecular potential rather than the use of empirical one—a point we have stressed for over a decade—is likely the correct strategy today, certainly the correct strategy tomorrow. It is gratifying to note that our strategy is becoming more and more accepted and that "ab initio" potentials are more and more used either as the sole "starting material" or in conjunction with empirical data.

From Low To High Relative Humidities

Likely the most classic physico-chemical laboratory experiments concerning DNA hydration are those pioneered by Lord and his school[57] and later extended by Vinograd[58] and others.[59] We have previously discussed in detail and compared some of these experimental findings with our computer experiments.[60] Below, we report additional data obtained from a B-DNA (G-C, A-T) sample containing 30 base-pairs and neutralized by 60 Na^+ ions.[42] The Monte Carlo experiments, carried out at a

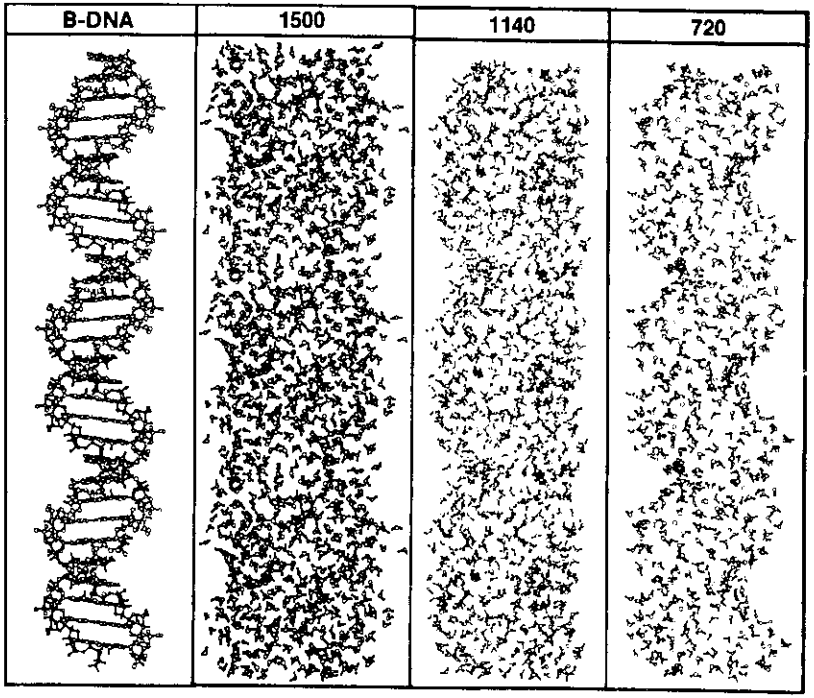
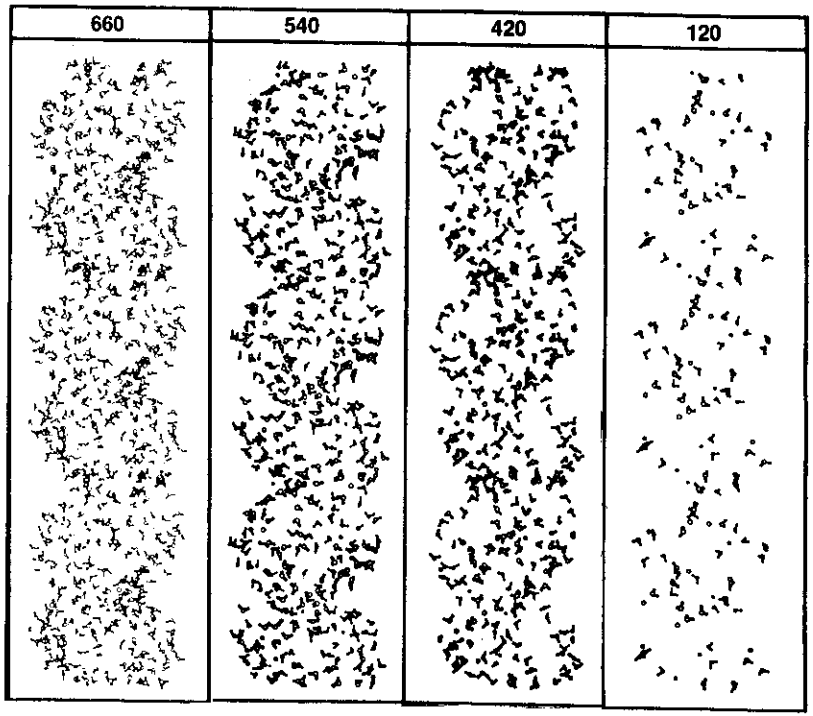


Figure 14. Patterns of water molecules and Na^+ counterions hydrating B-DNA from low to high relative humidities at 300K temperature. *Figure 14 continued on the next page.*

simulated temperature of 300°K are performed for a large range of relative humidity values and start with one water molecule per (neutralized) nucleotide unit,—or 60 water in total. The relative humidity is progressively increased up to 25 water molecules per nucleotide unit (or 1500 in total). For each relative humidity, a full Monte Carlo simulation is executed leading to predictions on the most probable position and orientation of the water molecules and Na^+ counterions as function of the relative humidity. The quantitative data are collected in a large number of tables, reported elsewhere;[61] here we summarize some of the results mainly using figures of immediate (but qualitative) value.

In Figure 14 we compare the water molecules and counterions orientation and positions for a statistically meaningful Monte Carlo conformation at relative humidities corresponding to 120, 420, 540, 660, 720, 1140, 1500 water molecules and 60 Na^+ counterions for our three turns B-DNA (G-C, A-T) fragment. *The collective nature of the interactions is most evident*: the B-DNA global attraction structures the ensemble of the water molecules and ions in such a way as to reproduce a global image of the B-DNA electric field. Note that we have not reported in each inset the B-DNA structure in order to simplify the drawing; the B-DNA structure is



reported only in the inset to the left of the one with 1500 water molecules but can easily be "inferred" from the insets. The image of the field becomes, however, less and less *visible* in the Figure when we increase the relative humidity above 700 water molecules, simply because more and more water are over-imposed on a very clear pattern still fully visible up to about 720 molecules. In Figure 15, a more quantitative view-point is given: in the inset at the left we show how many water molecules (ordinate axis) are either in the grooves, or in the first hydration shell (F.H.) or very strongly bound (b) to the DNA's atoms and counterions. The abscissae reports the number of water molecules, normalized to one B-DNA turn (thus, for example, the value 500 of Figure 15 corresponds to the value 1500 of Figure 14). The inset on the right reports the detailed partition of the hydration at different DNA sites. For example, when 300 water molecules hydrate one B-DNA turn, about 60 of these are in the grooves, 240 are in the first hydration shell and of these about 200 are strongly bound to B-DNA. More specifically (see inset on the right) between 3 and 4 water molecules hydrate each Na^+ ion, about 6 water molecules hydrates each (PO_4-CH_2) unit and about 10 water molecules hydrates either G-C or A-T. By a variation of the relative humidity, these values change as shown in the diagrams of Figure 15. Let us add few comments. *First*, the hydration number reported are obtained for those water molecules with the oxygen atoms within 3Å from he DNA atom (or ion or atomic groups) considered in the diagrams: by changing the value of 3Å even by only 0.5Å, the above hydration numbers change considerably, especially for high relative humidities. *This note is essential in order to properly compare* our data with laboratory experiments. *Second*, the above simulation results most nicely agree with Lord-type findings[57,58,59] as well with Dickerson's recent laboratory data, some obtained after our computer experiments were made available. *Third*, the structural data—only few are discussed in this article—are complemented by equally extensive and detailed energetic tabulations given elsewhere.[60-62] *Fourth*, details concerning major/versus minor grooves hydration are also discussed at length elsewhere.[60]

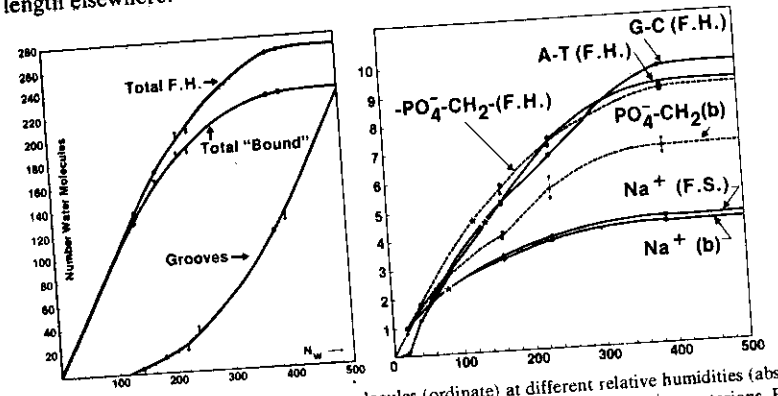


Figure 15. *Left*: Average number of water molecules (ordinate) at different relative humidities (abscissae) hydrating B-DNA or in the grooves. *Right*: same for specific groups and for Na^+ counterions. First hydration shell (F.H.) might contain water molecules not strongly bound (b.).

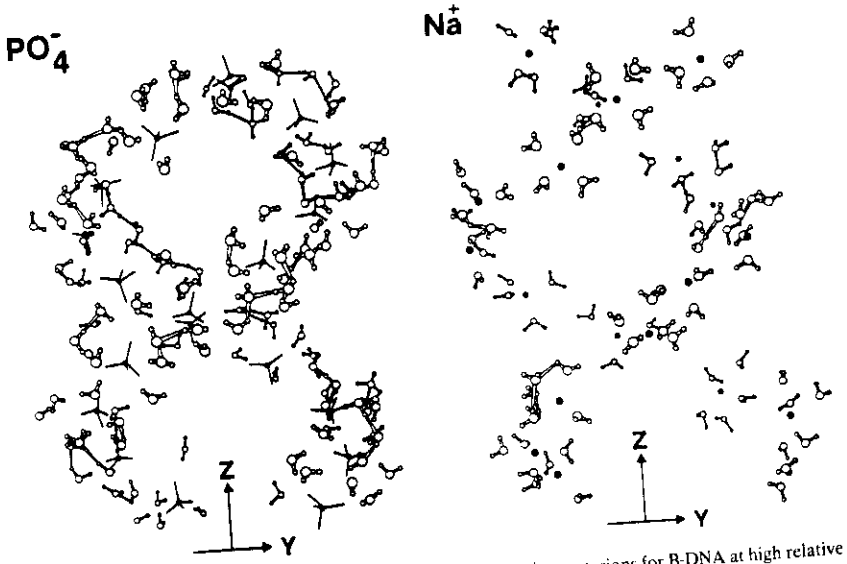


Figure 16. Patterns of hydration around the phosphates and Na^+ counterions for B-DNA at high relative humidities and 300K.

In Figure 16 we report a detail obtained from the 1500 water molecules Monte Carlo experiment; we have considered the hydration of the PO_4^- atoms (right inset) and Na^+ counterions. The PO_4^- are represented by a segment for each one of four PO bonds; the Na^+ ion, by a full dot placed at the position obtained by averaging many Monte Carlo conformations. The computed orientation of the water molecules is with the hydrogens atoms toward PO_4^- and away from Na^+. This is a *local* aspect in the hydration process. Bridging waters were found between two successive phospate, as first reported in our early computer experiments in the 1978-1979 period. Energetics details, complementing our structural determination have been made available elsewhere, and are therefore not discussed in detail in this work. As above noted, these computer experiments have been later confirmed by X-ray data; note, however, that strictly speaking, no water-water bridge has been experimentally detected up to now, since the X-ray resolution is much too low to allow for reliable measurements on the orientation of water molecules. However, the approximated determination of the position for some of the water molecules, allowed Dickerson et al. to guess the water molecules orientation,[3] and thus to discuss hydration patterns.

In Figure 17, we point out once more the *collective* and the *local* nature of the water molecules hydrating DNA. In the top inset we show molecules of water hydrating the base-pairs (the latter is shown simply as a segment; the asterisk added

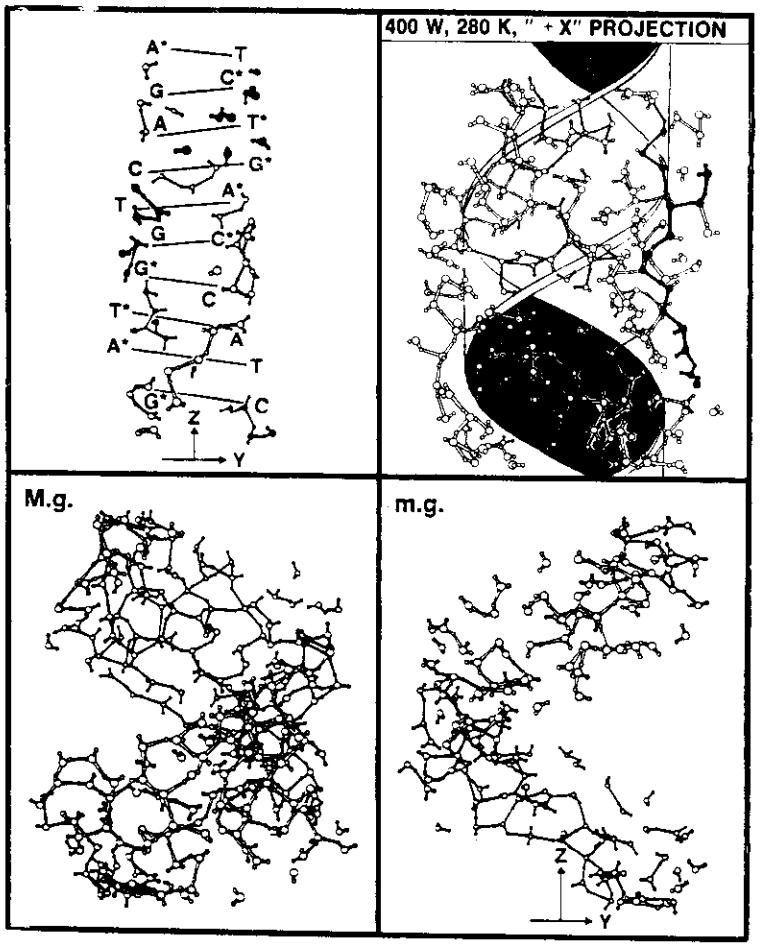


Figure 17. Patterns of hydration at the base-pairs in B-DNA (top, left) and at the major (top left and bottom right) and minor grooves.

to the base name indicates that the base belongs to one strand; base without asterisk designates those on the other strand). The specific position and orientation is related to the *local interaction* predominant near one atom of a base (or of the base-pair or of two-base pairs). There is, in addition, a very visible over-all pattern, forming two interwoven helices: this latter characterization constitutes the *collective aspect*. A very nice example of collective behavior is found also at the two grooves, where the water molecules assume two predominate patterns formed by filaments of water molecules one hydrogen bonded to the next one; one pattern is a

roto-translational pattern, and a second one is a pattern parallel to the Z-axis (the long axis of DNA). For details we refer elsewhere.[62,42,60,46] Again these findings seem to have been confirmed by recent X-ray experiments: if this is indeed the case, then it would add an interesting information, namely an indication of a rather extended lifetimes for some of the connectivity-pathways we have predicted. We are now in

In conclusion, in this section we have discussed both the *collective and the local nature* of the hydration as emerging from our computer experiments, corroborated by recent and preliminary experimental findings. The connectivity pathways fully envelope DNA and its counterions: the "structure" of DNA includes these hydrogen bonded structures of water molecules, which are as essential to a proper description of DNA as are the phosphate groups or the base-pairs or the counterions. Let us now turn our attention to the position of the counterions, which from Figure 16, appear to be arranged into regular patterns.

On The Structure And Dynamics Of Mono-Charged Counterions

From the iso-energy maps reported in the Figures 10, 11 and 12, one would expect to find counterions in the grooves near to the base-pairs and/or near the phosphate groups. The long range repulsion of the counterions, their mobility and the water-screening makes the above expectations somewhat more circumspect, but qualitatively it remains a very reasonable one. This is expected to be the case for either low or high relative humidities, and also for a solution, where, according to the results of condensation theory, the largest fraction of the counterions is located near DNA and only a small fraction is expected to be relatively farther away from DNA, inside the "solvent space". Unfortunately, condensation theory, models DNA very grossly, as an impermeable cylinder, thus it cannot differentiate between a phosphate or a base-pair, nor between B- or Z-DNA. In addition there are no explicit water molecules in the condensation theory, and one assumes the "same undifferentiated bulk water" either far from or near to DNA. On the other hand the agreement of condensation theory to some experimental data seems to indicate that a balanced cancellation of errors has been obtained. The Monte Carlo computer experiments have none of the limitation of either the iso-energy maps or condensation theory.

Even without including the sugar to base charge transfer, we did find[40,41,42,46] that Na^+ counterions penetrate the grooves and bring about—in the case of A- and B-DNA—a structure resembling two helices with different amplitude. This result—as far as we know—did constitute the first determination of a "structure" for the mono-charged counterions in B-DNA. Let us now consider in more detail our finding, referring also to new data[63] recently obtained.

First, we consider Na^+ counterions and B-DNA single helices with bases of only one kind, namely Poly-ASP, Poly-CSP, Poly-GSP and Poly-TSP. These four Monte Carlo simulations were performed with 1345, 1326, 1365 and 1410 water molecules, respectively and one Na^+ counterion per phosphate group, at a simulated tempera-

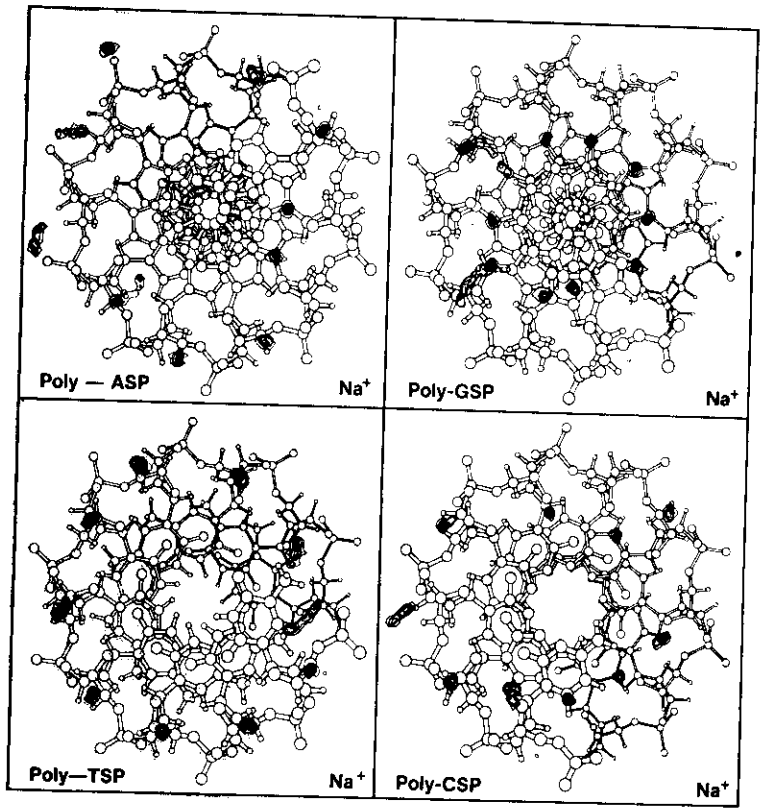
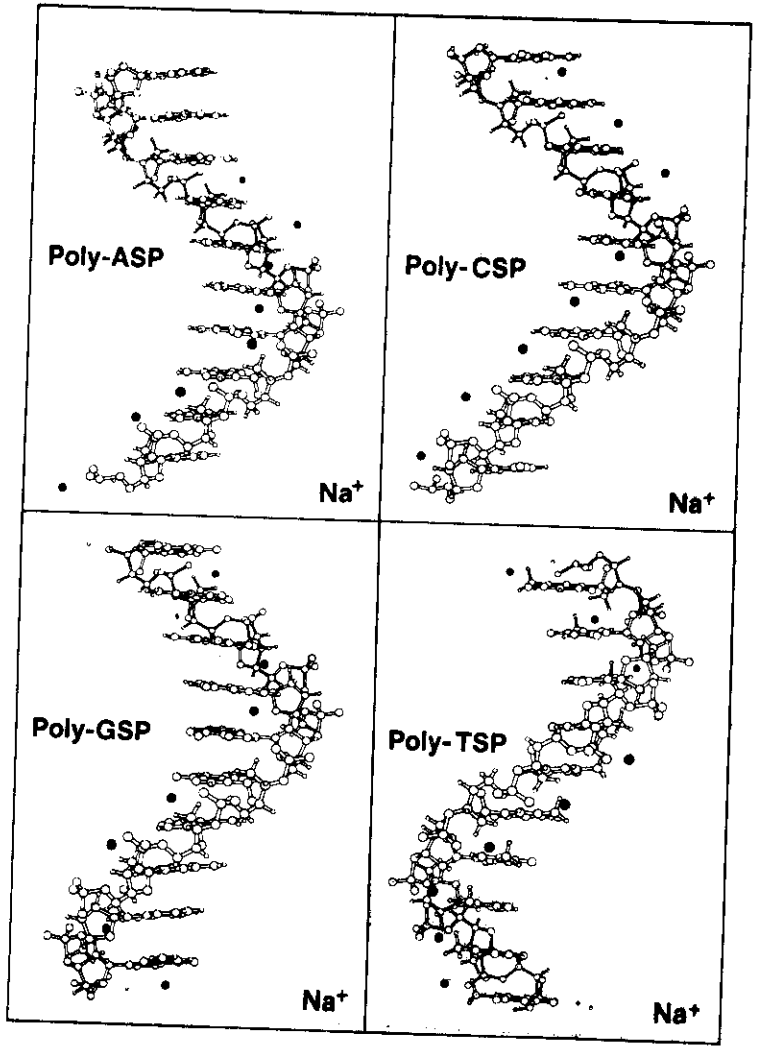


Figure 18. Na^+ counterions average position and iso-density maps for Poly-ASP, Poly-CSP, Poly-GSP and Poly-TSP shown in the x-y and x-z planes. *Figure 18 continued on the next page.*

ture of 300K. Three full turn fragments were considered, i.e., 30 bases in addition to the sugar and phosphate groups, SP. In Figure 18 we report the probability distributions (resembling spots of irregular shape) for the counterions and the projections of the B-DNA atoms either on the x-y plane or on the x-z plane. The counterions form an helix-like structure, with counterions positioned either near the bases (Poly-GSP) or the phosphates (Poly-TSP) or with some counterions near the bases and some near the phosphates (Poly-ASP and Poly-CSP). Therefore a G-C rich DNA and an A-T rich DNA will have not too different Na^+ distributions. Thus, these recent results concerning Na^+ counter-ions nicely confirm the double helix pattern of counterions previously found for the B-DNA double helix.[41] Notice that a rather extended region of space is characterized by high probability of containing the counterion, especially in the x-y plane (less in the z-direction). Notice in addition, that the mobility of a counterion (implied in the broad probability distribution) brings about mobility of the water molecules solvating the ion. Another



observation of dynamical type is that one can expect very highly *concerted motions* in all mechanisms where one counterion is displaced from his site. This was seen very nicely by following the evolution of the probability distribution maps obtained by placing counterions either at the boundary of the cylinder (containing the DNA and the water molecules) or at a very small R value. It should be noted that a concerted behavior is to be expected in a system characterized by long range interactions (like with ions); this system is expected to be particularly efficient for collective effects and for very fast transmissions along the z-direction. These obser-

vations should be relevant to unwinding, recognition and defect migration mechanisms. For more detailed information on the B-DNA single-helices interacting with Na^+, we refer elsewhere.[63]

Before concluding this section we wish to recall that the structure of a counterion at a polyelectrolite can be explained with the "polytopic bond" concept, presented for LiCN (where the DNA replaced the CN^- group previously studied[64]). This concept applied to our biological system has notable potentiality, and with its help, the dynamical characterizations of counterions become clearer. We recall that the probability distributions reported above (and in the next section) do not imply a localization of the ion at one site; this interpretation is supported by the finding that the number of high "probability sites" is larger than the number of the counterions, namely one ion has probability to be at *different sites*, as in polytopic bonds.

Specificity Of Counterions

Let us now return to the double helix, in particular to B-DNA (G-C, A-T), B-DNA (G-C, C-G) and Z-DNA. From many biological and biochemical processes we have become fully familiar with the expectation of ions *specificity*; ions carriers, enzymatic activity and membrane permeability are few among the many examples. Also in DNA ionic specificity is experimentally well established: one has only to recall that DNA undergoes conformational transitions by changing counterions;[46,59,60] the

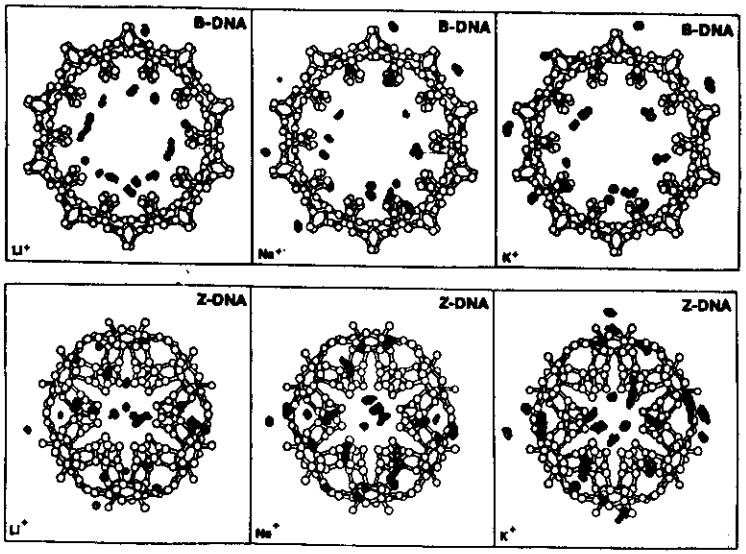


Figure 19. Probability distributions for counterions around B-DNA and Z-DNA after projection into the x-y plane.

iso-energy maps we have reported for B-DNA and for Z-DNA are certainly suggesting qualitatively different types of interactions.

In Figure 19 we report the probability distribution maps (see Ref. 6 on techniques to obtain such maps) comparing Li^+, Na^+ and K^+ in B-DNA and Z-DNA double helices. As by now, standard in our Monte Carlo simulations, we have selected to work with three full turns of DNA, 500 water molecules per turn and as many counterions as the number of phosphate groups, at a simulated temperature of 300K. The different positions assumed by Li^+ relative to Na^+ (or K^+) are clearly visible in the figure; equally evident are the differences for the same ion in the two conformers, B-DNA and Z-DNA. The hydration patterns are therefore also different and detailed tabulations on such information are available elsewhere[62,64] and briefly recalled below. In Table II we report the results concerning the first hydration of groups of atoms at few sites either in B- or Z-DNA. Because of the ambiguities in defining the radius for the first hydration shell at the atoms in DNA, the analysis has been performed for three different values of radius R. The numbers of water molecules reported in the table for the three different distances are obtained in the following way. We center at a given atom (site) a sphere of radius R and we count how many water molecules fall within the sphere (the check on the distance is done only for the oxygen atom of H_2O) during the entire Monte Carlo simulation about

Table II
Number of water molecules in the first hydration shell of selected sites, for Li^+, Na^+ and K^+. Comparison between B- and Z-DNA at three different values of the radius.

	B-DNA								
	R=2.5			R=3.0			R=3.5		
Site	Li^+	Na^+	K^+	Li^+	Na^+	K^+	Li^+	Na^+	K^+
PO_4^-	0.81	0.81	0.78	5.89	5.78	6.01	7.71	7.35	7.32
A	1.61	1.49	1.05	4.24	3.69	1.91	6.39	5.79	4.34
C	1.48	1.19	1.01	3.87	4.05	2.16	5.81	5.24	4.19
G	1.57	1.66	1.23	6.05	4.48	3.25	8.66	7.47	5.73
T	0.01	–	–	2.38	2.12	1.00	4.38	3.69	1.77
A-T	1.71	1.50	1.05	5.86	4.86	2.78	8.74	7.91	5.71
G-C	3.00	2.84	2.23	8.34	7.13	5.31	11.37	9.96	8.09
N.U.	2.02	1.90	1.61	9.06	8.31	8.00	11.35	10.54	10.01
ion+N.U.	2.20	3.14	4.92	2.20	3.14	4.92	2.20	3.14	4.92
ion	3.51	4.32	5.86	9.41	9.62	10.90	11.40	11.21	12.00
	Z-DNA								
	R=2.5			R=3.0			R=3.5		
Site	Li^+	Na^+	K^+	Li^+	Na^+	K^+	Li^+	Na^+	K^+
PO_4^-	1.00	0.85	0.85	5.64	5.82	5.32	7.13	7.23	6.44
C	1.54	1.27	0.87	2.80	2.51	1.46	4.03	4.11	2.56
G	0.93	0.96	0.58	3.74	3.60	1.66	6.37	6.00	3.74
G-C	2.47	2.22	1.45	5.08	4.79	2.72	7.45	7.39	5.14
N.U.	2.09	1.76	1.51	8.11	7.89	6.60	10.16	10.40	8.73
ion+N.U.	2.49	2.93	4.27	2.49	2.93	4.27	2.49	2.93	4.27
ion	4.10	4.32	5.26	8.54	8.79	8.80	10.56	10.57	10.17

10^6 configurations). As expected, by increasing the sphere radius, the number of water molecules falling within the volume increases proportionally. By reporting our result at three R values, we have covered the range of interest in most laboratory experiments. From the table we can see that for a given R, for example R=3.0Å, the number of water molecules belonging to the first hydration shell of a given group, notably changes from Li^+ to Na^+ to K^+. A finding is particularly interesting: it appears to be a common feature for the bases (A,C,G and T) and for the base pairs (A-T and G-C) to "loose" water if Li^+ is substituted with Na^+, or Na^+ with K^+, whereas the PO_4^- are solvated by roughly the same number of water molecules, independently from the kind of counterions in solution. This finding is equivalent to state that changing from one ion to another we find dehydration around the bases, whereas the phosphates are only partially affected. This finding is present both for B and Z-DNA. Another interesting finding is that in Z-DNA the solvation numbers of the water molecules at the phosphate groups are less constant than in B-DNA, in particular for the Na^+ and K^+ counterions.

As expected, the number of water molecules solvating the ions increases with the size of the ion. If we compare these values with those obtained for the same ions in solution (namely Li^+=4.0, Na^+=5.4, K^+=6.8) we can notice that the ions appear to be less solvated when in presence of DNA. In our opinion, however, the coordination around each ion remain about constant; in our computer experiment, some of the ions are often close to some of the atom of DNA and the coordination number should be complemented by including these atoms.

Statistical Mechanics of Midy-Systems: Extended Monte Carlo

Before introducing some of our latest results we have to briefly return on a problem of methodology. In the previous sections we have discussed Monte Carlo simulations where a fragment of DNA is enclosed in a cylinder containing the counterions and 1 to 25 water molecules per nucleotide unit. The cylinder has a radius R limited to values between 15 to 20 Å. From the statistical mechanical representation (and from Monte Carlo method in particular) we obtain predictions relevant to either high or low relative humidities or to DNA's first hydration layer in a solution. Our statistical representation does not allow—however—to verify aspects where bulk water need to be considered or where one asks for problems related to very large values of R. To discuss large R values, like about 100Å, one would need to consider about 1.5×10^5 water molecules, too many for practical considerations. (Indeed, computers have limits both in speed and in the random memory.) As an alternative we could select to work within the standard representations traditionally used to study polyelectrolites in solution, as for example, shown in Record's work[16] or in Manning's early papers on condensation theory (see references reported in Ref. 16). This alternative, however, is somewhat unpleasant, since the standard thermodynamical representation of the polyelectrolite structure is forced to be a very crude one. Indeed no distinction between A-, B-, Z-DNA is possible, since no phosphate-sugar-base characterization is retained; further, in the standard condensation theory, DNA is simplified to a uniformaly charged cylinder of parameterized

radius. In addition, we would not learn any detail about hydration patterns, either local or collective, since water molecules do not appear explicitly in the thermodynamical model. In other words, the very nature of the thermodynamical simulation prevents to obtain those details at the molecular and atomic level, we standardly obtain from the statistical mechanics representation and which are basic to a description at atomic resolution i.e., "chemical resolution."

The dilemma, either a very detailed description but for a miny-system (our DNA fragment, ions and at most few thousand water molecules), or a very general but undetailed description of a macro-system (for example, a DNA very grossly approximated, enclosed in an ionic solution without an explicit description of water molecules) can be eliminated by an approach which is rather new—as far as we know—and has been briefly summarized by us elswhere.[6] In this new approach—called "extended Monte Carlo"—we study a midy-system composed by DNA, few hundreds (or few thousands) water molecules, the counterions and a variable number of "solvated ions". The DNA fragment and the water molecules are kept inside a cylinder, called "inner volume", of radius R (15Å in our case), which is co-axial to a much larger cylinder called "outer volume" of radius R′ (100Å in our case). In the cylindrical volume with r such that $R' \geq r \geq R$, we allow "solvated ions". The two volumes are schematically represented in Figure 20. In our procedure the water molecules cannot leave the inner volume and are subjected to the standard Monte

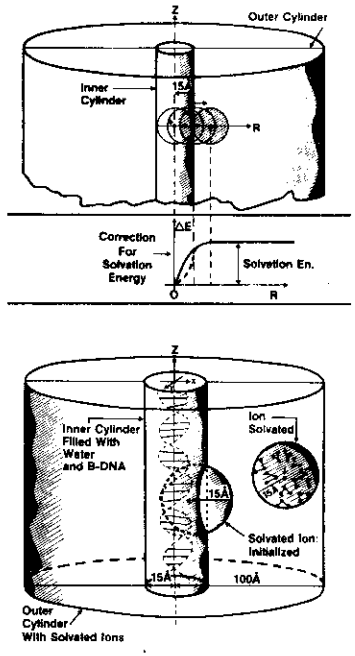


Figure 20. Extended Monte Carlo: Inner cylinder and outer cylinder for the inner and outer volumes; solvated ions and solvation correction.

Carlo sampling. For the ions we distinguish two cases: when inside or outside the inner volume. When the ion is within the inner volume, we assume that it experiences the interactions of each counterion in the outer volume and the particles within the inner volume (namely the standard Monte Carlo procedure discussed in the previous experiments is applied) and an additional interaction with those molecules of water assumed to be within a sphere of radius R(i) centered on the ion *but laying outside the inner volume*. The latter interaction is computed once for always (table look-up) by assuming it can be obtained from an analysis of the ion-water interaction energies computed as a function of the radial distance for a system composed by the water molecules filling up a spherical volume and with the ion at the sphere's center. When the ion is in the outer volume, it is assumed to be at the center of a solvation sphere of radius R(i). Two cases can occur: either the sphere is totally outside the inner volume or the sphere is partially outside the inner volume. In the former case we talk of "unperturbed-solvated ion", in the latter of "perturbed-solvated ion". An unperturbed solvated ion experiences the interaction energies of all the other unperturbed solvated ions (see below) and of the particles within the inner volume. A perturbed solvated ion experiences the interaction with all the other ions, with DNA and with the water molecules which are inside the inner volume; in addition a correction which accounts for the fraction of the solvation sphere which is outside the inner volume is added. Notice that the unperturbed-solvated ion, unperturbed-solvated ion interaction is the one we have previously discussed (see the section "Sites for Counterions and Solvated Ions Interactions") which can be precomputed once for always and tabulated as a function of the ion-ion separation.

In Figure 20 we schematically show the inner volume and the outer volume (top inset) and an ion in its solvation sphere at various distances from the inner volume boundary: the energy correction to be added is schematically represented in the central inset of the figure. The bottom inset shows an unperturbed solvated ion and a second ion with a solvation sphere partly inside the inner volume and partly outside, thus requiring an energetic correction as discussed above. In conclusion, in our model we retain *a fully detailed description, where it is needed*, namely in the inner volume, i.e., near DNA, which is modeled at the atomic level resolution; water molecules and ions are also explicitly described, when in the inner volume. Outside it an unrefined description is sufficient and the Monte Carlo technique is simplified by considering only solvated ions experiencing both other solvated ions in the outer volume and the water molecules, the counterions and DNA all in the inner volume.

DNA In Solution With Counterions

The *extended Monte Carlo* procedure has been applied to study the B-DNA fragment with three full turns (that is, with 30 base-pairs), 1500 water molecules in the inner volume and 60 K^+ counterions. The "extended Monte Carlo" technique did allow to determine how many counterions are in the inner or in the outer volume. In addition, since standard probability distribution maps can be constructed as in

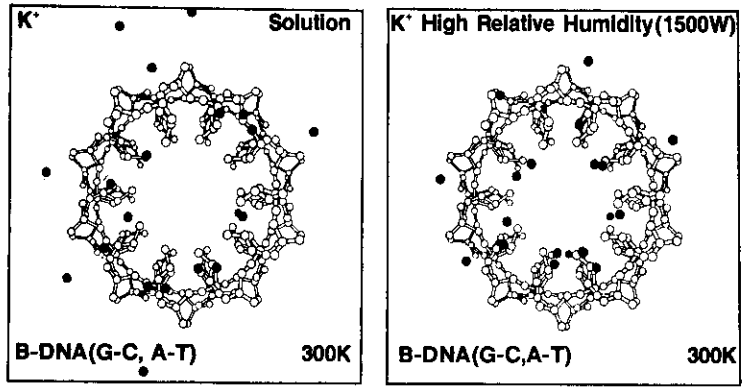


Figure 21. Average positions of K^+ counterions in solution (extended Monte Carlo) and at high relative humidity (standard Monte Carlo).

the previously reported Monte Carlo simulations, one obtains the probabilistic distribution of the cations in the outer and inner volume. In Figure 21 we show *average* positions for the K^+ when we experiment with the micro-system (standard

Table III

Comparison between standard Monte Carlo (left side) and extended Monte Carlo (right side). The first column reports the solvation number for each ion, the second the code for the nearest neighbor atom in DNA; the third column, the distance in Å from that atom.

ion	#H_2O	l.n.n.	r	#H_2O	l.n.n.	r
1	6.56	O3′	2.51	4.22	—	a)
2	4.11	O1′	2.69	6.00	O1′	3.83
3	5.26	O1P	2.48	4.00	O1P	4.34
4	6.44	O3′	3.35	6.03	O3′	4.35
5	5.37	O1′	2.46	5.00	N3	2.82
6	5.70	O1′	2.57	5.71	O1′	2.76
7	3.00	N3	2.45	6.91	C5′	3.52
8	7.51	C4′	3.44	9.00	—	a)
9	3.06	O2	2.77	3.00	O2	2.92
10	6.80	O3′	3.26	6.44	O3′	4.26
11	4.91	N7	2.54	5.11	N7	2.52
12	4.97	N7	2.57	7.08	O2P	4.61
13	4.05	O4	2.38	7.59	—	a)
14	4.88	O2P	2.69	5.91	O2P	2.60
15	4.91	N7	3.31	6.39	N7	2.57
16	6.42	O4	2.49	7.00	O4	3.90
17	7.39	N4	3.59	7.66	O2P	4.67
18	6.00	N7	2.66	4.00	—	a)
19	4.31	N7	2.47	7.29	N7	3.10
20	5.27	O4	2.56	7.27	C5M	3.92
Average	5.35		2.86	6.43		3.54

Monte Carlo) and when we experiment with the midy-system (extended Monte Carlo). We have represented each counterion with a dot representing its *average position* rather than with its probability distribution, since we wish to explicitly show each one of the 20 K^+ ions (per one B-DNA turn). Three ions are found, in average, outside the inner volume, the remaining are inside; in addition, one of the ions in the inner volume is far from any atom of DNA. Preliminary results are also reported in Table III, where we give an index (1 to 20) for each counterion, the number of water molecules solvating it, the counterions first nearest neighbor atom on the DNA fragment (if the ion is within 5 Å from a DNA's atom) and the corresponding distance (r value) obtained from the standard and from the extended Monte Carlo procedures. It should be noted that in Manning's condensation theory about 25% of the mono-charged counterions are "far" away from the polyelectrolite and the remaining are "condensed" at the polyelectrolite. From our experiment we obtained a distribution very close to the one predicted from the condensation theory, since four ions out of twenty are "far" away from the polyelectrolite. The condensed counterions, 80% of the total population, are all within the inner volume and for these the full details (at the atomic level) of the standard Monte Carlo modelings are available. We are now extending the above preliminary analyses to Li^+ and Na^+ counterions, having tested that our extended Monte Carlo technique and computer program is operational. The results reported in Table III, seem to confirm that Manning's condensation theory, even if based on gross approximations, can be assumed to be reliable at least for general predictions concerned with the overall distribution of the counterions.

Concluding Remarks

In this work two interests have been stressed: a methodological interest aimed at realistic simulations of complex systems and a chemical one on the interaction of solvent water molecules and a macromolecule. Concerning the first aspect, we have exposed the main lines of our "generalized model" which operationally links quantum chemistry to statistical mechanics to thermodynamics. Concerning the second aspect, we have stressed both the local and the collective nature of the solvent molecules, when interacting with DNA.

Our predictions on the counterions structure are expected to become of basic importance for an understanding of the DNA's dynamics, reaction mechanisms and base-pair recognition. The "extended Monte Carlo technique" allows us to state that our conclusions are valid both for high and low relative humidity and also for solutions. The existence of the water molecules connectivity pathways bring about the need to further extend our modeling by including simulations on the life-time of such structures and the related problem on the motions and modes of DNA. Further, the connectivity pathways, being confined to a very limited region of the phase-space, point to the need of an explicit account of entropic variations.

One more conclusion is firmly established: DNA in a solution with its counterions well defined structural patterns and with the water molecules connectivity path-

ways is *a very different bio-polymer* from the one where solvent molecules and counterions are neglected!

Acknowledgment

It is a pleasure to thank Drs. G. Corongiu and P. Otto for the use of unpublished data developed at our laboratory.[63,65]

References and Footnotes

1. Shakked, Z., Rabinovich, D., Cruse, W. B. T., Egert, E., Kennard, O., Sala, G., Salisbury S. A. and Viswamitra, M. A., *Proc. Roy. Soc., (London) B213*, 479 (1981).
2. Connor, B. N., Takano, T., Tanaka, S., Itakura, K. and Dickerson, R. E., *Nature, 295* (1982).
3. Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V. and Kopka, M. L., *Science, 216*, 475 (1982).
4. Wang, A. H. J., Fujii, S., van Boom, J. H. and Rich, A., *Proc. Natl. Acad. Sci., (USA) 79*, 3968 (1982).
5. Drew, H. R., Samson, S. and Dickerson, R. E., *Proc. Natl. Acad. Sci., (USA) 79*, 4040 (1982).
6. Clementi, E., "Computational Aspects for Large Chemical Systems", Lecture Notes in Chemistry, Vol. 19, Springer-Verlag, Berlin, Heidelberg, New York (1980).
7. Finney, J. L. and Goodfellow, J. M., this volume.
8. Beveridge, D., this volume.
9. Clementi, E., *Phys. of Elec. and Atomic Coll., VII ICPEAC*, North-Holland, 399 (1971).
10. Barker, J. A. and Watts, R. O., *Chem. Phys. Letters, 3*, 144 (1969).
11. Stillinger, F. H. and Rahman, A., *J. Chem. Phys., 60*, 1545 (1974).
12. Mehl, J. W. and Clementi, E., IBM Technical Report RJ #883, (June 22, 1971).
13. Lowdin, P. O., this volume.
14. Careri, G., this volume.
15. Palma, M. U., this volume.
16. Anderson, C. F. and Record, Jr., M. T., this volume. See in addition references given in Manning, G. S., this volume.
17. Lewin, S., *J. Theor. Biol., 17*, 181 (1967).
18. Rupprecht, A. and Forslind, B., *Biochem. Biophys. Acta., 204*, 304 (1970).
19. Bolis, G. and Clementi, E., *Chem. Phys. Letters, 82*, 147 (1981).
20. Clementi, E. and Popkie, H., *J. Chem. Phys., 57*, 1077 (1972); Kiskenmacher, H., Popkie, H. and Clementi, E., *J. Chem. Phys., 58*, 1689 (1973); ibid. *59*, 5842 (1973); ibid. *61*, 799 (1974).
21. Clementi, E. and Barsotti, R., *Theor. Chim. Acta., 43*, 101 (1976)
22. Clementi, E. and Barsotti, R., *Chem. Phys. Letters, 59*, 21 (1978).
23. Scheraga, H. A., this volume.
24. Dickerson, R. E., this volume; see in addition References 3 and 5.
25. Corongiu, G. and Clementi, E., *Gazz. Chim. Ital., 108*, 687 (1978); Clementi, E., Corongiu, G. and Lelj, F., *J. Chem. Phys., 70*, 3726 (1979).
26. Metropolis, N., Rosenbluth, A. W., Teller, A. H. and Teller, E., *J. Chem. Phys., 21*, 1078 (1953).
27. Clementi, E., "Determination of Liquid Water Structure: Coordination Numbers for Ions and Solvation for Biological Molecules", Lecture Notes in Chemistry, Vol 2, Springer-Verlag, Berlin, Heidelberg, New York (1976).
28. Alder, B. M. and Wainwright, T. W., *J. Chem. Phys., 31*, 459 (1959).
29. See for example some of the references given in Reference 16.
30. Pratt, L. and Chandler, D., *J. Chem. Phys., 67*, 3683 (1977); ibid. *73*, 3434 (1980).
31. Kiskenmacher, H., Popkie, H. and Clementi, E., *J. Chem. Phys., 58*, 1689 (1973).
32. Scrocco, E. and Tomasi, J., *Topics in Current Chemistry, 42*, 95 (1973).
33. Ranghino, G. and Clementi, E., *Gazz. Chim. Ital. 108*, 157 (1978).
34. Clementi, E. and Corongiu, G., *J. Chem. Phys., 72*, 3979 (1980).
35. a) Clementi, E. and Corongiu, G., *Proceeding of the Int. Symp. in Quantum Biology and Quantum*

Pharmacology, 9, P.O. Lowdin, Ed., pp.213-221, Wiley Intersciences, N. Y. (1982); IBM Research Report POK-09 (March 3, 1982); b) Otto, P., Clementi, E. and Ladik, J., *J. Chem. Phys.*, (submitted), IBM Research Report POK-13 (August 18, 1982).

36. Ladik, J. and Suhai, S., *Int. J. Quant. Chem., QBS7*, 181 (1980).
37. Ladik, J., this volume.
38. Lavery, R. and Pullman, B., *Nucleic Acid Research, 18*, 4677 (1981) and reference therein given.
39. Pullman, B., Proc. Symp. Steric Effects in Biomolecules, Eger, (Hungary) p. 237 (1981) and references therein given.
40. Pullman, A., Proc. Symp. Steric Effects in Biomolecules, Eger, (Hungary) p. 247 (1981) and references therein given.
41. Clementi, E., *IBM Journal of Research and Development, 25*, 315 (1981).
42. Clementi, E. and Corongiu, G., *Biopolymers, 21*, 763 (1982).
43. Clementi, E., Corongiu, G., Gratarola, M., Habitz, P., Lupo, C., Otto, P. and Vercauteren, D., *Int. J. Quant. Chem.*, (in press), *IBM Research Report POK-10* (March 8, 1982).
44. Sundaralingam, M. and Westhof, E., this volume.
45. Clementi, E. and Corongiu, G., *Biopolymers, 18*, 2431 (1979); *Int. J. Quantum. Chem., 16*, 897 (1979); *Chem. Phys. Letters, 60*, 175 (1979).
46. Clementi, E. and Corongiu, G., *Biomolecular Stereodynamics*, Sarma, R., Ed., Adenine Press, New York, p. 209-259 (1981).
47. Clementi, E., IBM Research Report POK-16 (1982). In this report we have added a number of figures either not included in this paper or reported with low detail.
48. Scordamaglia, R., Cavallone, F. and Clementi, E., *J. Am. Chem. Soc., 99*, 5545 (1977).
49. Corongiu, G. and Clementi, E., *Gazz. Chim. It., 108*, 273 (1978).
50. Rapaport, D. C. and Scheraga, H. A., *J. Chem. Phys. 86*, 873 (1982) and references therein given.
51. Clementi, E., Kolos, W., Lie, G. C. and Ranghino, G., *Int. J. Quant. Chem., 17*, 377 (1980); Habitz, P., Bagus, P., Siegbahn, P. and Clementi, E., *Int. J. Quant. Chemistry*, (in press); IBM Research Report POK-12 (May 10, 1982).
52. Clementi, E., Kistenmacher, H., Kolos, W. and Romano, S., *Theor. Chim. Acta., 55*, 257 (1980).
53. Kress, J. W., Clementi, E., Kozak, J. J. and Schwartz, M. E., *J. Chem. Phys., 63*, 3907 (1975).
54. Craven, B. M. and Mascarenhas, Y. *Acta. Cryst., 1Z*, 407 (1964).
55. Craven, B. M. and Takey, W. J., *Acta Cryst., 1Z*, 415 (1964).
56. Onuchic, J. N., Corongiu, G. and Clementi, E., (to be published).
57. Falk, M., Hartman, K. A. and Lord, R. C., *J. Am. Chem. Soc., 84*, 3843 (1962); ibid. *85*, 387 (1963); ibid. *85*, 391 (1963)
58. Hearst, J. E. and Vinograd, J., *Proc. Natn. Acad. Sci. (SA) 47*, 825 (1961); ibid. *47*, 999 (1961); ibid. *47*, 1005 (1961).
59. Texter, J. *Prog. Biophys. Molec. Biol., 33*, 83 (1978).
60. Clementi, E. and Corongiu, G., *Annals of the New York Academy of Sciences, 376*, 83 (1981).
61. Clementi, E. and Corongiu, G., *Int. J. Quantum Chemistry, 22*, 595 (1982).
62. Clementi, E. and Corongiu, G., *Biopolymers, 20*, 551 (1981); ibid. *20*, 2427 (1981).
63. Otto, P., Clementi, E. and Corongiu, G., (to be published), see also IBM Research Report POK-17.
64. Clementi, E., *J. Chem. Phys., 58*, 2460 (1973).
65. Clementi, E. and Corongiu, G., (to be published).
66. The data reported in this paper have been presented at the Sanibel meeting in honor of Prof. J. Mayer (March, 1982) exception made to those concerning the extended Monte Carlo simulations, which however, have been reported to the La Jolla International Symposium, "Structure and Dynamics of Nucleic Acids and Proteins" (September, 1982). This paper follows rather closely the format adopted for the 1982 Burke Memorial Lectures (Faraday Division of the Chemical Society).

Structure of Aggregates of Water and Li^+, Na^+, or K^+ Counterions With Nucleic Acid in Solution

ENRICO CLEMENTI and GIORGINA CORONGIU
IBM Corporation, P. O. Box 390, Poughkeepsie, New York

ABSTRACT: The position and orientation of water molecules hydrating fragments of DNA in the B and Z conformations are analyzed with the help of computer simulations. Monte Carlo studies are carried out at room temperature, high relative humidity (500 water molecules per pitch) and in the presence of counterions such as Li^+, Na^+, and K^+. Differences in hydration patterns and in the counterionic structures were found by compairing B-DNA with Z-DNA double helices and B-DNA helices with different base-pair distributions. The present extension of our similations to Z-DNA and to Li^+ and K^+ counterions permits some general conclusions concerning nucleic acids in solution.

INTRODUCTION

THE BASIC MACROSCOPIC CHARACTERIZATION OF DNA at the electronic and electrostatic level is that it is a polyelectrolyte. This fact, well appreciated since the early experiments on DNA, was the starting point for our statistical mechanical computer simulations as well as for some thermodynamic studies on DNA of the last 10 to 15 years. It is equally well known that the early DNA fibers diffraction studies were at such low resolution that much of the sugars and the bases, and all the DNA hydrogen atoms and water molecules and counterions, remained undetected. The neglect of the counterions had a limited impact on DNA gross structural determinations, which were based on X-ray data and unconcerned with energetics. Clearly, an equivalent neglect is totally unacceptable in any theorectical model dealing with either the electronic structure of DNA, i.e. the electronic states and bands, or the electrostatic potential and field. Unfortunately, counterions have been, and even today are often neglected in quantum chemical studies dealing with the above properties. This is a rather serious situation; indeed the electronic characterization of DNA without counterions is no more acceptable than a characterization of the structure of an atom without its nuclear charge! Whereas in an ionic crystal the position of the anions depend on the positions of the cations, in a rather equivalent way the positions of the phosphate groups in DNA are dependent on the positions of the counterions; however, the sugar geometry constrains the possible conformations of the phosphates and hence, indirectly, those of the counterions.

COMPUTATIONAL TECHNIQUES

The computational method used in this work has been proposed and tested in many publications over the last ten years. Therefore, we shall only recall its main features and refer extensively to the literature. In general a computation can be decomposed into three successive steps (summarized below).

Some of the data reported in this paper have been presented at the Sanibel meeting in honor of Prof. J. Mayer (March, 1982); the remaining data have been reported at the La Jolla International Symposium, "Structure and Dynamics of Nucleic Acids and Proteins" (September 1982) and in the 1982 Bourke Memorial Lectures—Chemical Society—Faraday Society Division, London, U.K

At first, we compute (by *ab initio* quantum mechanical methods) the interactions between two molecules, say A and B, one kept rigidly at a given position in space, the second placed at many positions and orientations, such as to reasonably sample the interaction potential hyperspace. In our case A can be a water molecule or a counterion and B another water molecule, a counterion, or a fragment of DNA. Depending on the nature of the intermolecular interaction energy, correlation effects are included either at the C. I. (configuration interaction) level or as a perturbation or, if very small, neglected. As is well known, the main electronic correlation correction in intermolecular forces is for dispersion effects. As is equally well known, when the intermolecular interaction is mainly of ionic type, the Hartree-Fock method is adequate, if we settle for a 5% to 10% error for the interaction energy. Often Hartree-Fock type basis sets are computationally too expensive, especially for molecules with more than 25-35 atoms, and in these cases we use a minimal basis set, m.b.s. (we never use "subminimal" basis sets. A minimal basis set is one which reproduces single zeta Slater-type energy for the isolated atoms; inferior-quality basis sets are referred to as "subminimal". Specifications on the basis sets we have used are contained in the papers dealing with atom-atom potentials). The large basis set superposition error, often present in m.b.s., is corrected with the well known counterpoise technique (C. P.) (Boys and Bernardi, 1970). A rather short, overall and very general presentation of this approach is given by Clementi (1980a); extensive discussions and tests are also given by Clementi (1976, 1980b). The importance of the combined use of m.b.s., C.P., and dispersion corrections has been analyzed in detail and tested by Kolos (1979), Kolos, et al. (1980a, 1980b), and Bolis, et al. (1983).

In the second step, we use the computed intermolecular interaction energies to construct (by fitting techniques) two-body atom-atom potentials, one atom belonging to the molecule A and the second to B. The atom-atom potentials are expressed in some analytical form, which can be either very simple (for example of the 12, 6, 1 type) or rather complex, especially if we wish to obtain a particularly accurate fit (with standard deviation smaller than 5%). Notice that the typical standard deviation of our fits is about 5% to 15%. A rather short and general discussion on the analytical forms on atom-atom potentials is given by Clementi, et al. (1981); extended discussions are available (Clementi, 1976, 1980b). Tabulations and analyses of specific atom-atom potentials are given for (ions-water Clementi and Popkie 1972; Kistenmacher, et al., 1973a, 1973b, 1973c, 1974; Popkie, et al., 1973; Clementi and Corongiu, 1978; Clementi, et al., 1980) for water-biomolecules. Bolis, et al., 1977; Carozzo, et al., 1978; Clementi, et al., 1977, 1979a, 1979b; Corongiu and Clementi, 1978; Corongiu, et al., 1979a; Ragazzi, et al., 1979; Scordamaglia, et al., 1977. For ions-biomolecules (Corongiu and Clementi, 1983; Corongiu, et al., 1979b, 1980), and for water-water (Kistenmacher, et. al., 1974a, 1974b; Popkie, et al., 1973). When more than two molecules interact, the use of pairwise potentials to obtain the

total intermolecular interaction energy is only an approximation; its three- and many-body corrections should be added for accurate studies. For example, at the energy minimum the three-body correction for the water trimer is about 1.5 kcal/mol, to be compared with the total two-body interaction energy for the trimer, which amounts to about 16 kcal/mol. The three-body correction (or pair-wise nonadditivity) in chemical systems containing water molecules and solutes has received detailed analyses and illustration (Clementi, 1980b; Clementi, et al., 1980a, 1980b; Habitz, et al., 1983; Kress, et al., 1975).

The third step is the use of the above analytical potentials as input for statistical mechanical simulations of Monte Carlo type, following the Metropolis algorithm (Metropolis, et al., 1953). In this step temperature and probabilistic considerations are introduced; as is known, both are neglected in the *ab initio* computations, i. e. in steps one and two. It is well known from the references already cited, that reasonably accurate atom-atom potentials. and Monte Carlo techniques can yield structural information in good agreement with X-ray and neutron beam experimental data. It is also known, however, that some thermodynamic data may be reproduced poorly unless three-body corrections are included; hence in the rest of this paper we shall use the Monte Carlo technique mainly to deal with structural predictions. Details on our computer programs and on specific Monte Carlo computations are available elsewhere for liquid water, (Lie and Clementi, 1975; Lie, et al., 1976) for water solutions containing ions (Clementi and Barsotti, 1976, 1978; Fromm, et al., 1975; Romano and Clementi, 1980; Watts, et al., 1974.) and for chemical systems composed of water molecules, counterions, and nucleic acids (Clementi, 1981, 1983; Clementi and Corongiu, 1979a, 1979b, 1979c, 1980, 1981a, 1981b, 1981c, 1982a, 1982b, 1982c; Clementi, et al., 1982.) have discussed the sites of water binding at nucleic acids. Very approximately, in DNA, all the hydration sites that are highly stable and permit the hydrogen atoms of the water to point toward DNA are in principle good candidates to be counterion-binding sites also. This similarity is obvious on the basis of electrostatic considerations. However, the range of the ion-ion interaction deserves some consideration; whereas a water-water interaction is nearly zero at an oxygen-oxygen distance of 9-10 Å, two counterions strongly repel each other at these distances. Ionic hydration cuts down the repulsive interaction; water molecules solvating an ion not only add stabilization to the system (the ion-water interaction) but also decrease the ion-ion repulsion. This screening is very important for understanding the role of water association with DNA and can be nicely simulated with Monte Carlo computer experiments, as recently demonstrated (Clementi, 1983).

Indeed solvated ion--solvated ion interactions can be computed, for example, by considering many water molecules and two ions, for example two K^+, placed inside a sphere of radius R; the two ions are positioned symmetrically displaced from the center of the sphere by an amount L, where $L \ll R$. In the recent computer procedures (Clementi, 1983) R. is 15 Å and $L \leq 5$ Å, so that the ions are always far away from the boundaries of the sphere. For very large separation of the two ions, we switch to a different Monte Carlo technique, considering two spheres of radius $R' > R$, each with an ion at its center; R' is such as to make the sphere contain more molecules of water than the computed coordination number. The computed interaction (Clementi, 1976; Clementi and Barsotti, 1976, 1978; Fromm, et al., 1975; Romano and Clementi, 1980; Watts, et al., 1974) energies, at short and long ion-ion distances, are then used to obtain the fitting constants for an analytical expression of the solvated ion–solvated ion interaction. As expected, the solvated ion repulsions are weaker than the bare ion repulsions.

The decrease in the electrostatic repulsion due to water molecules located between two cations is general for ionic charges separated by a dielectric; therefore, equivalent decreases are present for the repulsion of any two phosphate groups on nucleic acid chains and separated by water molecules.

These general considerations show a few of the fundamental aspects in the hydration of DNA. As just noted, adding water to DNA decreases the phosphate-phosphate repulsion by screening, and therefore stabilizes the DNA system. At the same time counterions condense around DNA, further decreasing the phosphate-phosphate repulsion and bringing about electrostatic stabilization. Second, water molecules solvate the counterions, thus decreasing their mutual repulsion, which would have destabilized the system. Further, since ionic interactions are very long range, one expects to encounter in the DNA + water + counterions system very prominent collective effects and concerted motions for the counterions and the molecules of water. Indeed, this has been verified in computer studies elsewhere discussed (Clementi, 1983; Clementi and Corongiu, 1982; Clementi, et al., 1982). Finally, since the interaction energies of cations with water and DNA are very specific (for each cation), specific and local effects will complement the collective effects mentioned. As shown elsewhere (Clementi and Corongiu, 1982a), these effects are essential in explaining, for example, the ion-selective transition from one DNA conformation to another.

It follows that any discussion and modeling of nucleic acid stability (and therefore, the energy balance for nucleic acid conformational transitions) must consider both interactions between negatively charged units (namely, the phosphate groups and also the bases, which are partially negatively charged because of electronic charge transfer (Clementi and Corongiu, 1982b) and also interactions of positively charged groups (the counterions and the sugar units, the latter partially positively charged because of charge transfer, (Clementi and Corongiu, 1982b).

INTERACTIONS OF COUNTERIONS, WATER, AND DNA SITES

ISO-ENERGY MAPS FOR COMPONENTS OF DNA

Iso-energy maps can be instructive for some understanding of the interaction between water and DNA (or its components) or between ions and DNA. As a prerequisite one needs to be in position to compute, reliably and very rapidly, the interaction energy between either a molecule of water or an ion and DNA. As is known (Clementi, 1976, 1980b) this prerequisite is met with the availability of atom-atom interaction potentials, already mentioned.

In the last several years this graphical technique has been also used to display electrostatic potential maps, following the proposal of Scrocco and Tomasi (1973). Such maps are by now very popular because of the ease of obtaining them. It should be stressed, however, that these maps, which are necessarily limited to representing the electrostatic interaction of a point charge with the electron density of a molecule, neglect many effects, such as polarization, charge transfer, dispersion, and exchange contributions to the total interaction energy. As a consequence, for example, the electrostatic approximation does not permit one to recognize specificity for the interactions of Li^+, Na^+, and K^+ with DNA; further, the electrostatic approximation fails when the point charge marginally overlaps with the electronic density distribution of the solute and completely breaks down when the charge substantially overlaps the density distribution, –, in the above example, near equilibrium and at shorter distances.

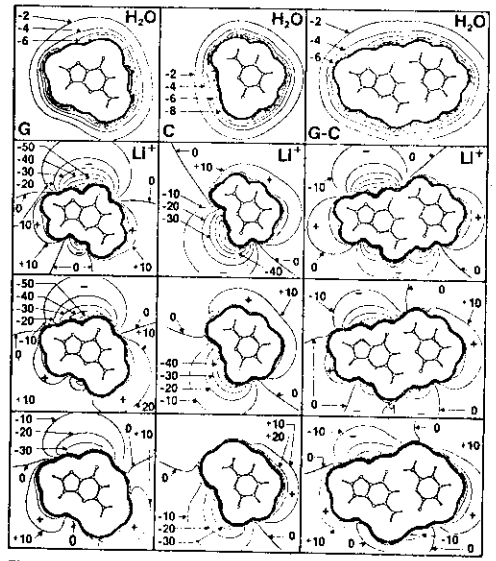


Figure 1. Iso-energy maps for G (left), C (center) and G-C pair (right) interacting with water (top line), Li^+ (second line), Na^+ (third line), and K^+ (bottom, or fourth line of insets).

Figure 1 shows iso-energy maps for two separated bases guanine, G, and cytosine, C, and for the base-pair, G-C (left, center, and right column, respectively). The iso-energy maps are computed for a plane containing the molecular plane. The first line of maps in Figure 1 reports iso-energy contours for one molecule of water; the second, third, and fourth (bottom) lines report maps of interactions with a Li^+, a Na^+, and a K^+ion, respectively. The iso-energy maps for G, C and G-C in Figure 1 (top line) are graphically superior to those previously reported (Clementi and Corongiu, 1980). In these maps the contour-to-contour energy difference is 2.0 kcal/mol for the interactions with water, and 10.0 kcal/mol for the interactions with the ions; repulsive contours above a given threshold have been omitted. One can easily see the attractive regions at the borders of the hard core. To help visualization, we have explicitly indicated the attractive (negative sign) and the repulsive regions (positive sign).

The strongly repulsive regions, which start at the zero kcal/mol contour and rapidly increase in value as one moves toward the inside of the molecule, limit and define the molecular hard core. As one can verify by inspection, the shape of the hard core is different when obtained by considering the interactions with a water molecule rather than with a counterion. Had we computed the C, G, and C-G iso-energy maps not for interactions with water, but with some other neutral molecule, again the shape would have been somewhat different. The same conclusion is reached by considering the interaction of our bases and/or base-pairs with either X-rays or neutron beams. Sometimes it is forgotten that the structure of a molecule is different for each different probe. It is also very obvious that a substituted G and/or C (such as a methyl or an ethyl homolog) would bring about very different maps.

The attractive and repulsive regions are qualitatively different in comparing the maps for water with those for the three ions. In addition, quantitative differences are seen in comparing the map for any one of the three ions with that for another ion. Because the insets are rather small, we point out that besides the easily visible differences, the size of the molecular core of G, C, and G-C decreases from K^+ to Na^+ to Li^+ by an amount proportional to the corresponding ionic radii.

One more point is of interest: the base-pair G-C attraction to the ions occurs at positions corresponding to the major and minor grooves in DNA, whereas the repulsion for ions occurs at two sides of the base-pair in correspondence to positions occupied by the sugar units. Thus, the base-pair seems particularly tailor-made to interact with ions when connected to the sugar-phosphate groups. But this is not all! As reported elsewhere, (Clementi and Corongiu, 1982b), the sugar unit in DNA donates about one quarter of an electronic charge to its base: thus each base is negatively charged and the G-C base-pair has, in total, about one half of an electron in excess acquired by the transfer. Obviously, because of elementary electrostatic considerations, this charge transfer notably increases the attraction of a counterion to the bases and base-pairs and at the same time makes the base-base interaction somewhat weaker, because of the charge transfer-induced electrostatic repulsion. The computations of the charge transfer in B-DNA and Z-DNA were done by considering an oligomer with three sugar and two phosphate units and two bases, G and C, in a conformation corresponding to either B- or Z-DNA (Clementi and Corongiu, 1982b). Because of the large size of the above fragments, these computations correspond to the most extensive *ab initio* computations ever performed in quantum chemistry and quantum biology. The charge transfer has also been evaluated in *ab initio* computation of the electronic bands in B-DNA (Ladik, 1983; Ladik and Suhai, 1980; Otto, et al., 1982) using crystal orbital techniques. However, in the latter calculations, numerical complexity, forced the introduction of simplifications not needed in the molecular computation on the B-DNA and Z-DNA fragments. This point is duscussed in detail by Ladik (1983). Parenthetically, our overall electrostatic model is notably different from those in which both the counterions at the bases and the charge transfer from the sugar units are neglected; this difference in modeling widens considerably after our inclusion of solvent effects, nonzero temperature, and statistical averaging.

In previous papers (Clementi, 1981, 1983; Clementi and and Corongiu, 1981; Clementi, et al., 1982a) we pointed out that the strong attraction between the counterion and the base-pair is not limited to a counterion located in the base-pair molecular plane, but is present also when the counterion is above that plane. In this geometrical configuration the counterion might penetrate the base-pair in the hydrogen bond region and influence the base-to-base separations and orientations. This is tantamount to stating that the counterion can actively influence the base-pair opening mechanisms, the relative angular orientation of the two bases and the distances between them--all the vital parameters of nucleic acid con-

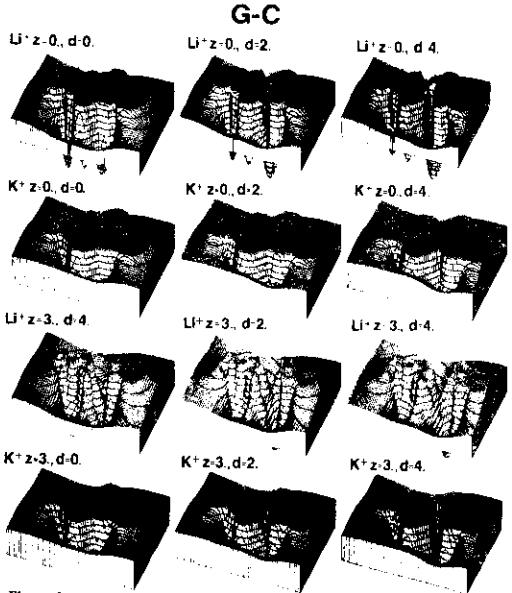


Figure 2. Three-dimensional maps of G-C with Li^+ and K^+ at z = 0 a.u. (top and third rows), z = 3 a.u. (second from top and bottom rows) for d = 0 a.u. (left column), d = 2 a.u. (center), and d = 4 a.u. (right).

formations. In Figure 2 we consider three-dimensional iso-energy maps for Li^+ and K^+ interacting with G-C either in the G-C molecular plane (z = 0 a. u. first and second line of insets) or above it by 3.0 a. u. (z = 3 third and bottom lines of insets). In addition, the two bases are either at the standard (equilibrium) base-pair geometry (d = 0.0 a. u. or are analyzed after increasing the distance between them by either 2.0 a. u. or 4.0 a.u. (d=2.0, d=4.0, respectively)(1.00 a.u.=0.529 Å). The two-dimensional maps for G-C in Figure 1, specifically the right-side insets on the second and fourth (bottom) lines, correspond to the two insets of Figure 2 designated as Li^+, z = 0, and K^+, z = 0., d = 0.. In Figure 2 the horizontal lines are iso-energy contours. We can see (look at the deep "canyon" in between the two bases, top line, right insert) that Li^+ can penetrate between the two bases even when it is constrained to stay in the molecular plane as long as the bases are pushed apart a bit (z = 0.0 a. u. and d = 4.0 a. u.). For K^+ this does not occur, because K^+ is less attracted to the bases and is larger. If we constrain Li^+ and K^+ in a plane 3.0 a. u. above the base-pair, both ions are attracted when immediately above the hydrogen-bonding region and d = 4.0 a. u. For Li^+ this attraction remains even at d = 2.0 a. u. and d = 0.0 a. u. Notice that the field of the phospates favors the above processes, as one can easily deduce by inspecting the iso-energy curves reported, for example in Figure 70 of Clementi (1980b).

Thus we have shown that the counterion can penetrate very near or be in between the two bases of a base-pair, bringing about an attractive energy which balances the loss of energy due to breakage of the hydrogen bonds. It is difficult to reconcile this finding with the frequent neglect of considerations about the effects of counterions on intercalations, defects, kinks, mutations, base-pair stability, etc., -- the vital processes of DNA reactivity and dynamics.

What we have stated above for the G and C bases and the G-C pair could now be restated for A, T, and the A-T base-pair, as clearly shown by the insets of Figure 3. Notice however, the existence of quantitative differences between the G-C and A-T base-pairs. The interactions of A, T, and A-T with water are not discussed here, since they are available elsewhere, (Clementi and Corongiu, 1980).

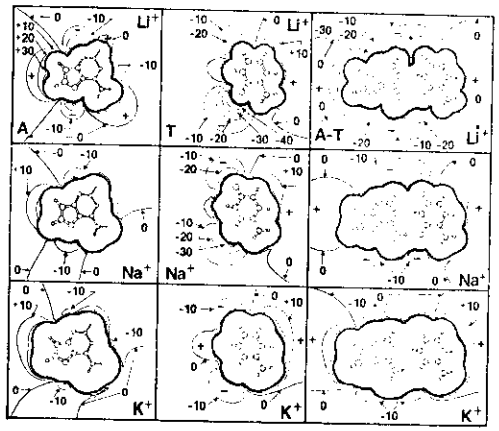


Figure 3. Iso-energy maps for A (left), T (center), and A-T base-pair (right) interacting with Li^+ (top), Na^+ (middle), and K^+ (bottom).

We conclude this section by recalling Sundaralingam's (1983) long-standing indentification of the central role played by the sugar unit in establishing the structural conformation for DNA: those experimentally based deductions are here complemented by our theoretical modeling, where the charge transfer from the sugar to the base is seen as a primary process, notably enhancing the attraction of counterions to the bases and/or base-pairs and weakening the base-pair interaction. Thus the critical role of the sugar unit for controlling conformation emerges also at the electronic structure level. There are two main positions that produce constraints for the double helix phosphate groups, one at the sugar units, and the second at the base pairs; the charge transfer weakens the base-base attraction, and thus the constraint at the base-pair, and adds flexibility to the polymer.

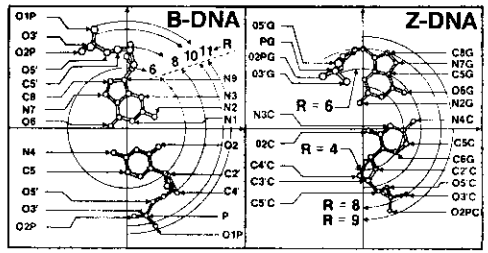
ISO-ENERGY MAPS OF THE B- AND Z-DNA - WATER SYSTEMS

A few years ago we reported preliminary Monte Carlo simulations of the interactions of a few water molecules with A- and B-DNA and the corresponding iso-energy maps (Clementi and Corongiu, 1979a, 1979b, 1979c). In those early studies, the atom-atom pair potentials did not include the charge transfer from the sugar to the base units already discussed. Below we analyze a few iso-energy maps for the interaction of a water molecule with B- and Z-DNA, including the effect of the charge transfer. The maps of Figures 4 and 5 were obtained by flattening out the iso-energy map corresponding to a surface of a cylinder of radius R, coaxial to the DNA long axis. The maps were computed for a fragment of DNA composed of three full DNA turns (thus the fragment contains 30 base-pairs in the case of B-DNA); only the central pitch is fully reported in the Figures. To help visualization each map is shown twice, once in two dimensions (right) and once in three dimensions (left). In the two dimensional representation, the ordinate corresponds to the length of one full turn (33.8 Å for B-DNA and 44.58 Å for Z-DNA) and the abscissa corresponds to 2-R, where R is the value selected for the cylindrical map circumference. In Table I, we report for a few R-values the atoms bisected by the cylindrical surface (or very near to it) for B-DNA and Z-DNA; these atoms constitute the hard core shown in the maps. As shown in Figure 1, the core repulsion energies are reported only up to the threshold value, 10 kcal/mol: this brings about the "mesa"-like appearance of the hard-core regions. Clearly, in B-DNA the smaller the value of R, the larger the hard-core region and correspondingly the smaller the grooves region (Major and minor groves are designated as M.g. and m.g., respectively). This is not true for Z-DNA, as seen from the data reported in Table I. We note that the designation B-DNA (G-C; C-G) is introduced for a B-DNA sample with base-pairs G-C and C-G; the designation B-DNA (G-C, A-T) is used to denote a B-DNA sample with 50% G-C and 50% A-T, elsewhere discussed (Clementi and Corongiu, 1981a, 1982a). In previous papers only the B-DNA (G-C; A-T) sample was considered so that the above differentiation in the designation was unnecessary.

For B-DNA (Figure 4) the M.g. valley floor is at about -15 kcal/mol for R = 6 Å; the values of -20 and -25 kcal/mol are found only near the edges of the valley, immediately before the beginning of the hard-core ridges. The m.g. floor is energetically more attractive to water as expected, much more narrow. These findings, available for some time (Clementi and Corongiu, 1979a, 1979b, 1979c), have been lately indirectly confirmed by experiment (Connor, et al., 1982; Dickerson, et al., 1982; Drew, et al., 1982; Wang, et al., 1982). For R = 8 Å we note in B-DNA the expected decrease in the width of the hard-core "mesa",

Table I. Atoms Participating in the Hard Core in the Cylindrical Maps

B-DNA (G-C, C-G)	
R = 6	O1', C1'
R = 8	O3', P, O2P, O5', C5', C4', C3', C2'
R = 10	O1P
R = 11	O1P
Z-DNA	
R = 4	O1'C, C3'C, C2'C, C4C, C5C, C6C, O3'G, N3G, C2G, N1G, C6G
R = 6	O1PC, O5'C, C5'C, C4'C, N4C, PG, O1PG, O2PG, C5'G, C3'G, C4G, O6G, C5G
R = 8	O3'G, PC, O2PC, O5'C, C4'G, O1'G, C2'G, C1'G, N9G, C8G, N7G,
R = 9	O2PG



Atoms near cylindrical cross section at R = 6, 8, 10, and 11 Å in B-DNA and R = 4, 6, 8, and 9 Å in Z-DNA (see also Table I).

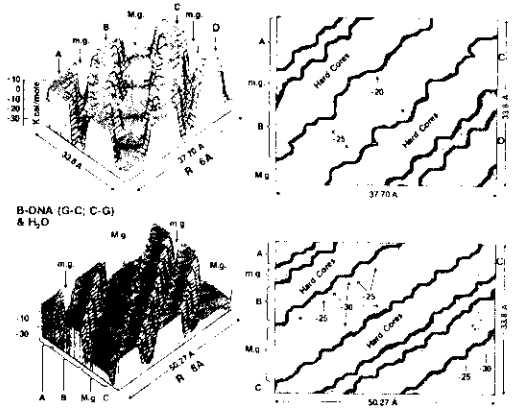


Figure 4. Cylindrical iso-energy maps for B-DNA with H_2O.

For Z-DNA the maps are very different, as evident from Figure 5. We observe that the one-groove map at R = 6 Å evolves, at R = 8 Å, into an "apparent" two-groove map with one groove deep and the second somewhat shallow; the latter corresponds to those atoms which are on the exterior of Z-DNA. Only the deeper groove survives at smaller R values. It is evident that such large differences in counterion-DNA interactions are bound to bring about notably different hydration patterns in B-DNA relative to Z-DNA and consequently, energetically different stabilizations in the two conformers.

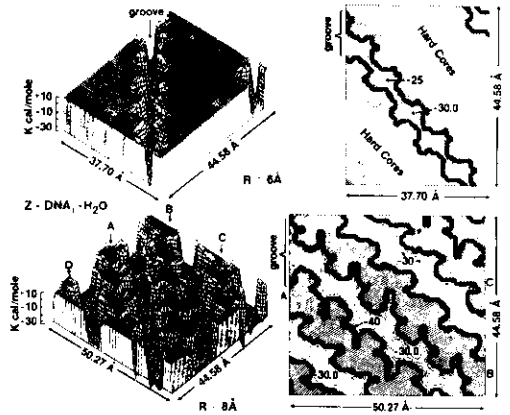


Figure 5. Cylindrical iso-energy maps for Z-DNA with H_2O.

ISO-ENERGY MAPS FOR LI^+, NA^+, K^+ WITH B- AND Z-DNA

Figure 6 displays the iso-energy maps for interactions of K^+ with B-DNA (G-C; C-G) at R = 6, 8, 10, and 11 Å. Starting at the larger R value (R = 11 Å) one sees the small-core regions corresponding to the oxygen atoms O1P at the phosphates

(see Table I), then a larger cross section of the individual phosphate groups (R = 10 Å), and then the large valley of m. g. and the larger M. g. valley (R = 8, 6 Å). Comparing the maps for R = 6 Å and R = 8 Å with corresponding one for water (Figure 8), shows that K^+ will have higher mobility than H_2O. The energy values reported at the valley floors in Figure 6 are extremely large. Clearly, they are totally unattainable in any laboratory experiment; DNA would not survive these energies, which are of the order those of plasma type. Indeed in no experimental condition could three full pitches of B-DNA be stable when only one counterion (or one positive charge) is present. For this reason the iso-energy maps offer only a preliminary and qualitative, though useful, picture; the same hold for the electrostatic or electric field maps.

A very interesting feature can be noted at R = 6 Å: the K^+ counterion can penetrate at the base and is nearly able to pass across grooves, for example from the major to the minor grooves. This observation is important because it implies that a counterion can come very close to the hydrogen bonds of the base-pairs and therefore interfere with the base-to-base bonding as previously discussed in reporting Figure 2. Note, in addition, the small differences in the energy at the valley floor from R = 6 Å to R = 11 Å; this is one of the main reasons for the extended mobility in the x-y plane (perpendicular to the main DNA axis) computed for the Na^+ counterion.

In Figure 7, we report the iso-energy maps for K^+ with Z-DNA at R = 4, 6, 8 and 9 Å. The existence of only one groove is evident at R = 4 and 6 Å; at R = 8 and 9 Å we approach the external atoms of Z-DNA and an apparent second groove develops, responsible for the many cusps shown in the hard-core regions at R = 6 and 4 Å.

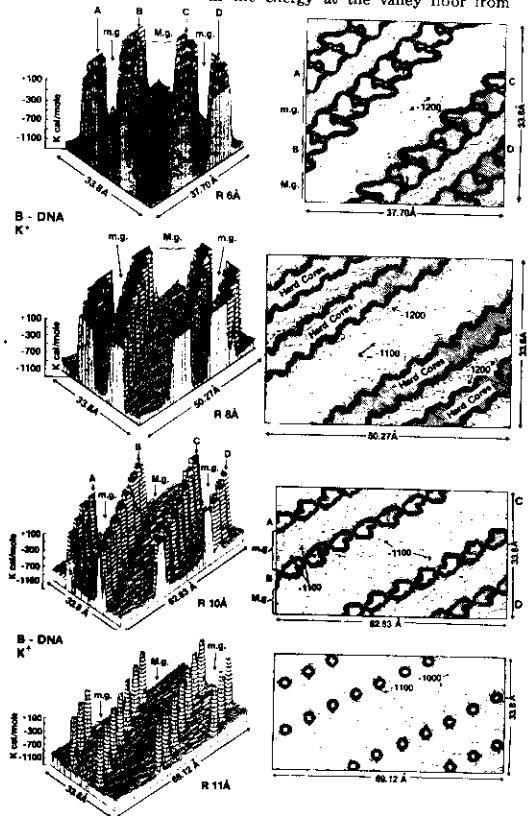


Figure 6. Cylindrical iso-energy maps for B-DNA with K^+ at R = 6, 8, 10, and 11 Å.

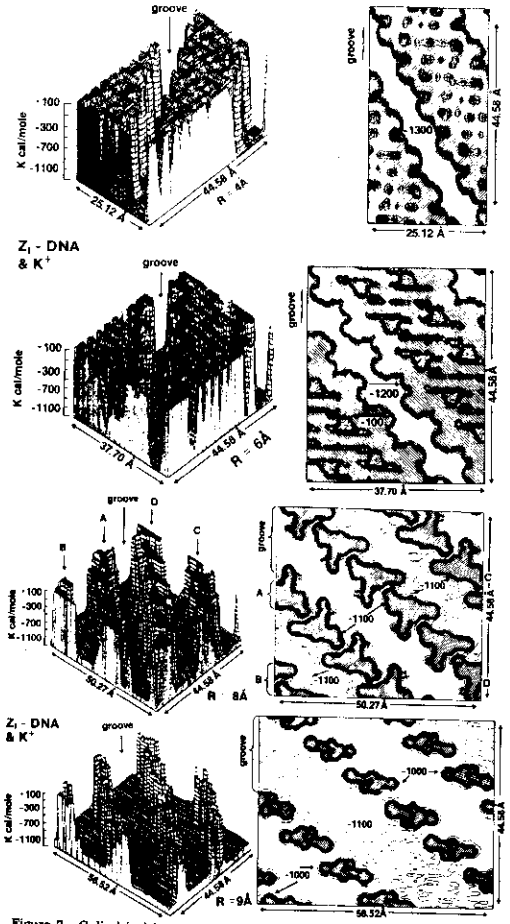


Figure 7. Cylindrical iso-energy maps for Z-DNA with K^+ at R = 4, 6, 8, and 9 Å.

Finally, Figure 8 compares for B-DNA (G-C; C-G) the iso-energy maps for Z-DNA and Li^+, Na^+ and K^+ at R = 6 Å. The valley floor energies become smaller (more attractive) in the order Li^+, Na^+, and K^+. Inspection of Figure 8 suggests that among the differences between the counterions one should include specific variations in the mobility.

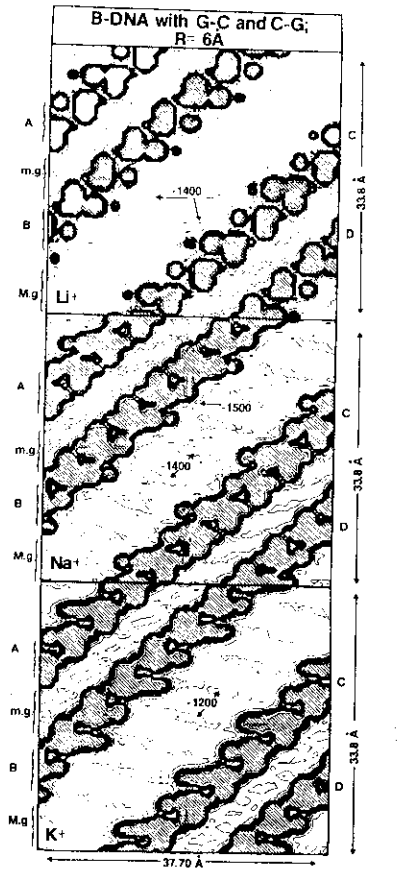


Figure 8. Cylindrical iso-energy maps for B-DNA and Li^+, Na^+, and K^+ at R = 6 Å.

DYNAMICS OF DISTRIBUTION OF SINGLY CHARGED COUNTERIONS

From the iso-energy maps in Figures 6, 7, and 8, one would expect to find counterions in the grooves near the base-pairs or near the phosphate groups. The long-range repulsion of the counterions, their mobility, and the water screening make the above expectation somewhat more circumspect, but qualitatively it remains a very reasonable one. This is so not only for low and high relative humidities, but also for solutions. According to condensation theory (Fixman, 1979; Manning, 1980; Manning and Zimm, 1965; Queran and Weisbuck, 1980; Record, 1967; Schellman and Stigler, 1977; Schildkraut and Lifson, 1965), the large fraction (about 80%) of the counterions is indeed located near DNA and only a small fraction is expected to be relatively farther away from DNA, inside the "solvent space". Unfortunately, condensation theory models DNA very grossly and therefore it cannot differentiate between a phosphate or a base-pair, nor between B- or Z-DNA. In addition, there is no explicit inclusion of the water molecules in condensation theory; one assumes the "same undifferentiated bulk water" either far from or near to DNA. (Other rather dubious approximations are accepted, but agreement with experiments indicates that, probably mainly because of cancellations of errors, condensation theory can lead to reasonable results). It is known that Monte Carlo computer experiments have none of the above limitations. We now report on our new Monte Carlo simulations obtained at high relative humidity (500 water molecules per pitch) and 300 K temperature.

Even without including the sugar-base charge transfer, we found (Clementi, 1981), that Na^+ counterions "penetrate" the groves and are structured, in the cases of A- and B-DNA, into two patterns resembling two helices with different amplitudes, one with the ions near the bases and one with the ions near the phosphates, each along a different groove. This result so far as we know, constituted the first determination of the arrangement of singly charged counterions in association with B-DNA. Our finding has been confirmed by new data (Otto, et al., 1983) recently obtained on single helices of B-DNA make up of bases of only one kind, namely poly-ASP, poly-CSP, poly GSP, and poly-TSP. Four Monte Carlo simulations (performed with 1345, 1326, 1365, and 1410 water molecules respectively and with one Na^+ counterion per phosphate group, at a simulated temperature of 300 K) again showed that the counterions form helix-like structures, with such counterions positioned either near the bases (poly-GSP) or the phosphates (poly—TSP), or with some counterions near the bases and some near the phosphates (poly-ASP and poly—CSP). Thus, these recent results nicely confirm the counterion helix patterns previously found for the B-DNA double helix (Clementi, 1983; Clementi and Corongiu, 1981a, 1982a, 1982b; Clementi, et al., 1982).

The above results indicate that a G-C-rich B-DNA will have counterion structural patterns not very different from those in an A-T- rich B-DNA. Below we shall return to this point. A rather extended region of space is characterized by high probability of containing the counterion, especially in the x-y plane (less in the -z-direction). The mobility of a counterion (implied in the broad probability distribution) brings about mobility of the water molecules solvating the ion. Another observation of dynamical type is that one can expect very highly concerted motions in all processes where one counterion is displaced from its site. This was seen very nicely by following the evolution of the probability distribution maps obtained by placing counterions either at the boundary of the cylinder (containing the DNA and the water molecules) or at a very small R value. A concerted behavior is to be expected in any system characterized by long-range interactions as between ions); such a system is expected to be particularly efficient for collective effects and for very fast transmissions along the z-direction. These observations should be relevant to mechanisms of unwinding, recognition, and defect migration.

The orientation of a counterion at a polyelectrolyte may be explained by use of the "polytopic bond" concept (Clementi, 1973) proposed long ago to account for some very peculiar features of LiCN molecules at high temperature; in the present study, DNA replaces the CN- group. In a polytopic bond, an ion such as Li^+ can orbit around a negatively charged fragment, such as CN-, at a very low energy cost. A standard structural characteristic, the position of Li^+ relative to CN-, is thus replaced by a characteristic orbit described by Li^+ around CN- at a given temperature. We speculate that the counterions might be constrained to the two helical-patterns discussed

above, but can very easily occupy nearly any position in the helix; clearly, if one ion moves from one position to another, all the counterions will also move in a concerted way.

We note that the probability distributions reported for DNA (either above or in the next section) do not imply a localization of the ion at one specific site. This is supported by the finding that the number of high-probility sites is larger than the number of counterions, so that one ion has high probability to be at different sites, as in polytopic bonds. Thus, counterions are very mobile.

SPECIFICITY OF COUNTERIONS

Let us now return to the double helix, in particular to B-DNA (G-C, A-T), B-DNA (G-C, C-G), and Z-DNA. From many biological and biochemical processes, we have become fully familiar with ionic specificity: this is manifested in ion carriers, enzymatic activity, and membrane permeability among many examples. Such specificity is experimentally well established in DNA also: one has only to recall that DNA undergoes conformational transitions when counterions are changed (Clementi and Corongiu, 1981a, 1981c; Drew, et al., 1980; Hanlon, et al., 1978; Ivanov, et al., 1974; Pohl, et al., 1972; Texter, 1978; Wang, et al., 1980; Zimmerman and Pheiffer, 1980). The iso-energy maps for B- and Z-DNA certainly suggest qualitatively different types of interactions.

Figure 9 shows probability distribution maps comparing Li^+, Na^+, and K^+ in B-DNA and Z-DNA double helices obtained by Monte Carlo simulations on systems with three full turns of DNA, 1500 water molecules, and as many counterions as phosphate groups (the simulated temperature is 300 K). The different positions assumed by Li^+ relative to Na^+ (or K^+) are clearly visible in the Figure; equally evident are the differences for the same ion in the two conformers, B-DNA and Z-DNA. Therefore, the hydration patterns are also different, as shown in Table II giving results on the first hydration of cations at some sites in either B- or Z-DNA. Because of the possibility of ambiguities in defining the radius R for the First hydration shell at the atoms of DNA, the analysis has been performed for three different values of R. To obtain numbers of water molecules shown in the Table for the three different distances, we center on a given atom (site) a sphere of radius R and count how many water molecules fall within the sphere (the distance is evaluated only for the oxygen atom of H_2O) during the entire Monte Carlo simulation (about 10^6 configurations). As expected, increasing the sphere radius proportionally increases the number

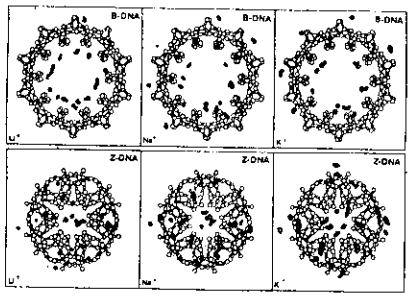


Figure 9. Probability distributions for counterions around B-DNA and Z-DNA after projection into the *x-y* plane.

Table II. Number of water molecules in the first hydration shell of selected sites, for Li^+, Na^+, and K^+. Comparison between B- and Z-DNA at three different values of the radius.

B-DNA (G-C; A-T)

	R = 2.5			R = 3.0			R = 3.5		
Site	Li^+	Na^+	K^+	Li^+	Na^+	K^+	Li^+	Na^+	K^+
PO_4^-	0.81	0.81	0.78	5.89	5.78	6.01	7.71	7.35	7.32
A	1.61	1.49	1.05	4.24	3.69	1.91	6.39	5.79	4.34
C	1.48	1.19	1.01	3.87	4.05	2.16	5.81	5.24	4.19
G	1.57	1.66	1.23	6.05	4.48	3.25	8.66	7.47	5.73
T	0.01	.....	.....	2.38	2.12	1.00	4.38	3.69	1.77
A-T	1.71	1.50	1.05	5.86	4.86	2.78	8.74	7.91	5.71
G-C	3.00	2.84	2.23	8.34	7.13	5.31	11.37	9.96	8.09
N.U. ion	2.02	1.90	1.61	9.06	8.31	8.00	11.35	10.54	10.01
N. U. + ion	2.20	3.14	4.92	2.20	3.14	4.92	2.20	3.14	4.92
	3.51	4.32	5.86	9.41	9.62	10.90	11.40	11.21	12.00

Z-DNA

	R = 2.5			R = 3.0			R = 3.5		
Site	Li^+	Na^+	K^+	Li^+	Na^+	K^+	Li^+	Na^+	K^+
PO_4^-	1.00	0.85	0.85	5.64	5.82	5.32	7.13	7.23	6.44
C	1.54	1.27	0.87	2.80	2.51	1.46	4.03	4.11	2.56
G	0.93	0.96	0.58	3.74	3.60	1.66	6.37	6.00	3.74
G-C	2.47	2.22	1.45	5.08	4.79	2.72	7.45	7.39	5.14
N. U.	2.09	1.76	1.51	8.11	7.89	6.60	10.16	10.40	8.73
ion	2.49	2.93	4.27	2.49	2.93	4.27	2.49	2.93	4.27
N. U. + ion	4.10	4.32	5.26	8.54	8.79	8.80	10.56	10.57	10.17

N. U. = necleotide unit.

or water molecules falling within the volume. Reporting results at three R values covers the range of interest in most laboratory experiments. Table II shows that for a given R, for example R = 3.0 Å, the number of water molecules in the first hydration shell of a given group changes noticeably from Li^+ to Na^+ to K^+. It is usual for the bases (A, C, G, and T) and for the base pairs (A-T and G-C) to "lose" water if Li^+ is replaced by Na^+, or Na^+ by K^+, whereas the phosphate groups are solvated by roughly the same number of water molecules independently of the kind of counterions in solution. This finding is equivalent to stating that changing from one ion to a larger one causes dehydration around the bases, whereas the phosphates are little affected. This is true for both B and Z-DNA. Also notable is that in Z-DNA the hydration numbers at the phosphate groups are more variable than in B-DNA, in particular for the Na^+ and K^+ counterions.

As expected, the number of water molecules solvating the ions increases with the size of the ion. Comparison of these values with those for the same ions in solution (Li^+ = 4.0, Na^+ = 5.4, K^+ = 6.8), shows that the ions are, on average, less solvated in the presence of DNA, since some of them can be very close to some of the atoms of DNA. For example, when a K^+ ion is near the phosphate groups, it is hydrated by 5 to 6 water molecules, whereas, when near the bases it is hydrated by only 3 to 4 water molecules.

The hydrations of B-DNA (G-C; C-G) and B-DNA (G-C; A-T) are not dissimilar. Our Monte Carlo simulation for B-DNA (G-C; C-G) at 300 K, 500 water molecules per pitch, and K^+ counterions, yields for the first hydration shell the following number of water molecules at the sites (in parentheses) previously studied at R = 3.5 Å (Table II): 6.82 (PO_4^-), 1.34(C), 2.05 (G); 2.73 (G-C), 8.04 (N.U.), and 4.02 (K^+); the major variation is at the base-pairs.

CONCLUSIONS

From these studies reported throughout this paper one major conclusion emerges: the single and double helices of DNA in biological solution should be regarded as "super-molecules" composed of the "old-fashioned X-ray - type DNA structure", counterions in definite patterns and water molecules with filaments of hydrogen bonds pathways we have often discussed.

The stability of this super-structure implies electroneutrality: the negative charges are located at the phosphate groups and at the base units (the latter are less prominent than the former, about one quarter of the total). The positive charges (the counterions) are located mainly very near DNA (at the phospate and at the base) and form helical skeletal shapes. Some of the counterions can also be found in the surrounding solute, a few water units away from DNA. From recent preliminary simulations we know that about one quarter of the counterions can be "dispersed" in the solvent rather than "condensed" around DNA. Thermodynamic, (Fixman, 1979; Manning, 1080; Manning and Zimm, 1965; Quiran and Weisbach, 1980; Record, 1967; Schellman and Stigter, 1977), and our statistical mechanical modeling, (Clementi, 1983), as well as experimental data (Connor, et al., 1982; Dickerson, et al., 1982; Drew, et al., 1982; Wang, et al., 1982) all seem to agree on this point.

The positions for the negative charges are heavily constrained; this brings about induced constraints for the positions of the positive charges. The positive charges are of two types: the primary positive charges (condensed counterions) and secondary positive charges (dispersed counterions and the charges at the sugar units). Equivalently there are primary negative charges (the phosphates) and secondary negative charges (the bases).

The negative charges 1) at the phosphates, 2) at the bases, and 3) the positive charges at the sugar unit are by definition structured in helical patterns. Thus, our finding of helical shapes for the primary positive charges is most reasonable.

Dynamical studies on DNA must account for the fact that the negative charges are constrained in a specific way characteristic for each DNA conformation, as dictated by the sugar geometry, whereas the positive charges are somewhat more free. Only about one quarter of the positive charges (those on the sugar units, which donate electronic charges to the bases) are as constrained as the negative charges. The mobility of the positive charges is, however, not as full as in the case of polytopic bonds, and one expects the counterion motions to be concerted. In addition, the solvent near DNA is made up of structured water molecules and the specificity of the connectivity pathways (Clementi, 1983; Clementi and Corongui, 1982c) is bound to affect the counterion dynamics (and vice versa). We recall that bulk water also is structured, but differently than the solvation water near DNA. Therefore, a counterion near or far from DNA will be affected differently by the different structural organizations of the two "types" of water.

We can also speculate that vibrational modes (Eyster and Prohovsky, 1974a, 1974b, 1977; Lindsay and Powell, 1983) present in the "old-fashioned" DNA structure will be perturbed by the counterion structures, (dumping-type perturbation) and by the hydrating water structures. On the other hand, Brownian motions in the "old-fashioned" DNA system might be enhanced by the counterions, because of the long-range action of ionic charges. Detailed computer simulations and laboratory experiments are needed to validate these dynamical hypotheses.

One point however, emerges clearly: our "super-molecule" DNA model is expected to be of basic importance in molecular recognition, in conformational transitions and, in general, in problems related to collective effects, (Connor, et al., 1982; Dickerson, et al., 1982; Drew, et al., 1982; Wang, et al., 1982) since it attempts to realistically account for the electrostatics of DNA behavior in solution. The "old-fashioned" DNA model is usually adequate as a first-order approximation, at least for some qualitative studies where solvent, counterions, temperature, and dynamics are ignored, as for some problems limited to local features in DNA. Finally, let us note that the "old-fashioned" DNA model recognizes only structures with long lifetimes, whereas the "super-molecule" DNA model is concerned also with structures with shorter lifetimes, for example, with evanescent ones that might last only as long as the time scale of a chemical reaction.

REFERENCES

Bolis, G.; Clementi, E., 1977. *J. Am. Chem. Soc., 99*, 5550.

Bolis, G.; Clementi, E.; Wertz, D. H.; Scheraga, H. A.; Tosi, C., 1983. *J. Am. Chem. Soc., 105*, 355.

Boys, S. F.; Bernardi, F., 1970. *Mol. Phys., 17*, 553.

Carozzo, L.; Corongiu, G.; Petrongolo, C.; Clementi, E., 1978. *J. Chem. Phys., 68*, 787.

Clementi, E., 1973. *J. Chem. Phys., 58*, 2460.

Clementi, E., 1976. *Determination of Liquid Water Structure: Coordination Numbers for Ions and Solvation for Biological Molecules, Lecture Notes in Chemistry*, Vol. 2. Springer-Verlag, Berlin, Heidelberg, New York.

Clementi, E., 1980b. *Computational Aspects for Large Chemical Systems, Lecture Notes in Chemistry*, Vol. 19. Springer-Verlag, Berlin, Heidelberg, New York.

Clementi, E., 1980a. *J. Phys. Chem., 84*, 2122.

Clementi, E., 1981. *IBM J. Res. Develop., 25*, 315.

Clementi, E., 1983. In *Structure and Dynamics: Nucleic Acids and Proteins*. Clementi, E.; Sarma, R. Eds., Adenine Press, N.Y., pp. 321-364.

Clementi, E.; Barsotti, R., 1976. *Theor. Chim. Acta., 43*, 101.

Clementi, E.; Barsotti, R., 1978. *Chem. Phys. Letters, 59*, 21.

Clementi, E.; Cavallone, F.; Scordamaglia, R., 1977. *J. Am. Chem. Soc., 99*, 5531.

Clementi, E., Corongiu, G., 1978. *J. Chem. Phys., 69*, 4885.

Clementi, E.; Corongiu, G., 1979a. *Biopolymers, 18*,2431.; 1979b, *Int. J. Quantum Chem., 16*, 897,; 1979c, *Chem. Phys. Lett., 60*, 175.

Clementi, E.; Corongiu, G., 1980. *J. Chem. Phys., 72*, 3979.

Clementi, E.; Corongiu, G., 1981. *Biopolymers, 20*, 551.

Clementi, E.; Corongiu, G.; Ranghino, G., 1981. *J. Chem. Phys., 74*, 578.

Clementi, E.; Corongiu, G., 1981a. *Ann. Acad. Sci., 376*, 83.

Clementi, E.; Corongiu, G., 1981b. *Biopolymers, 20*, 2427.

Clementi, E.; Corongiu, G., 1981c. In *Biomolecular Stereodynamics*, Sarma, R., Ed., Adenine Press, New York, pp. 209-259.

Clementi, E.; Corongiu, G., 1982a. *Biopolymers, 21*, 763.

Clementi, E.; Corongiu, G., 1982b. *Proc. Int. Symp. Quantum Biol.; Quantum Pharmacol., 9*, pp. 213-221. P. O. Lowdin, Ed., Wiley-Interscience, N. Y.,; IBM Research Report POK-09 (March 3, 1982).

Clementi, E.; Corongiu, G., 1982c. *Int. J. Quantum Chem., 22*, 595.

Clementi, E.; Corongiu, G.; Gratarola, M.; Habitz, P.; Lupo, C.; Otto, P.; Vercauteren, D., 1982. *Int. J. Quant. Chem. Symp., 16*, 409.; IBM Research Report POK-10 (March 8, 1982).

Clementi, E.; Corongiu, G.; Jonsson, B.; Romano, S., 1979. *FEBS Lett., 100*, 313.

Clementi, E.; Corongiu, G.; Jonsson, B.; Romano, S.; 1980. *J. Chem. Phys., 72*, 260.

Clementi, E.; Corongiu, G.; Lelj, F., 1979. *J. Chem. Phys., 70*, 3726.

Clementi, E.; Kistenmacher, H.; Kolos, W.; Romano, S., 1980. *Theor. Chim. Acta., 55*, 257.

Clementi, E.; Kolos, W.; Lie, G. C.; Ranghino, G., 1980. *Int. J. Quantum Chem., 17*, 377.

Clementi, E.; Popkie, H., 1972. *J. Chem. Phys., 57*, 1077.

Connor, B. N.; Takano, T.; Tanaka, S.; Itakura, K.; Dickerson, R. E., 1982.

Nature, 295.

Corongiu, G.; Clementi, E. (to be published).

Corongiu, G.; Clementi, E., 1978. *Gazz. Chim. Ital., 108,* 273.

Corongiu, G.; Clementi, E.; Dagnino, M.; Paoloni, G., 1979. *Chem. Phys., 40,* 439.

Corongiu, G.; Clementi, E.; Pretsch, E.; Simon, W., 1979. *J. Chem. Phys., 70,* 1266.

Corongiu, G.; Clementi, E.; Pretsch, E.; Simon, W., 1980. *J. Chem. Phys., 72,* 3096.

Dickerson, R. E.; Drew, H. R.,; Connor, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L., 1982. *Science, 216,* 475.

Drew, H. R.; Samson, S.; Dickerson, R. E., 1982. *Proc. Nat. Acad. Sci. USA, 79,* 4040.

Drew, H.; Takano, T.; Tanaka, S.; Itakura, K.; Dickerson, R. E., 1980. *Nature, 286,* 565.

Eyster, J. M.; Prohofsky, E. W., 1974. *Biopolymers, 13,* 2505.

Eyster, J. M.; Prohofsky, E. W., 1974. *Biopolymers, 13,* 2527.

Eyster, J. M.; Prohofsky, E. W., 1977. *Biopolymers, 16,* 965.

Fixman, M. J., 1979. *J. Chem. Phys., 70,* 4995.

Fromm, J.; Clementi, E.; Watt, R. O., 1975. *J. Chem. Phys., 62,* 1388.

Habitz, P.; Bagus, P.; Siegbahn, P.; Clementi, E., 1982. *Int. J. Quantum Chem.,* in press: see also same authors in IBM Research Report, POK-12, May 10, 1982.

Hanlon, S.; Chan, A.; Berman, S., 1978. *Biochim. Biophys. Acta., 519,* 526.

Ivanov, I.; Minchenkova, L. E.; Minyat, E. E.; Frant-Kamenetskii, M. D.; Schyolkina, A. K., 1974. *J. Mol. Biol., 87,* 817.

Kistenmacher, H.; Lie, G. C.; Popkie, H.; Clementi, E.; 1974. *J. Chem. Phys., 61,* 546.

Kistenmacher, H.; Popkie, H.; Clementi, E., 1973. *J. Chem. Phys., 58,* 1689.

Kistenmacher, H.; Popkie, H.; Clementi, E., 1973. *J. Chem. Phys., 58,* 5827.

Kistenmacher, H.; Popkie, H.; Clementi, E., 1973. *J. Chem. Phys., 59,* 5842.

Kistenmacher, H.; Popkie, H.; Clementi, E., 1974. *J. Chem. Phys., 61,* 799.

Kistenmacher, H.; Popkie, H.; Clementi, E.; Watts, R. O., 1974. *J. Chem. Phys., 60,* 4455.

Kolos, W., 1979. *Theor. Chim. Acta., 51,* 219.

Kolos, W.; Corongiu, G.; Clementi, E., 1980a. *Int. J. Quantum Chem., 17,* 775.

Kolos, W.; Ranghino, G.; Clementi, E.; Novaro, O., 1980b. *Int. J. Quantum Chem., 17,* 429.

Kress, J. W.; Clementi, E.; Kozak, J. J.; Schwarts, M. E., 1975. *J. Chem. Phys., 63,* 3907.

Ladik, J., 1983. In *Structure and Dynamics: Nucleic Acids and Proteins,* Clementi, E.; Sarma, R., Eds., Adenine Press, N. Y.

Ladik, J.; Suhai, S., 1980. *Int. J. Quantum Chem., QBS7,* 181.

Lie, G. C.; Clementi, E., 1975. *J. Chem. Phys., 62,* 2195.

Lie, G. C.; Clementi, E.; Yoshimine, M., 1976. *J. Chem. Phys., 64,* 2314.

See in addition Boys (1970).

Lindsay, S. M.; Powell, J., 1983. In *Structure and Dynamics of Nucleic Acids and Proteins,* Clementi, E.; Sarma, R., Eds., Adenine Press, N.Y., pp. 241-260.

Manning, G. S., 1980. *Biopolymers, 19,* 37.

Manning, G. S.; Zimm, B. H., 1965. *J. Chem. Phys., 43,* 4250.

Matsuoka, O.; Clementi, E.; Yoshimine, M., 1976. *J. Chem. Phys., 64,* 1351.

Metropolis, N.; Rosenbluth, A. W.; Teller, A. H.; Teller, E., 1953. *J. Chem. Phys., 21,* 1078.

Otto, P.; Clementi, E.; Corongiu, G. (to be published).

Otto, P.; Clementi, E.; Lakik, J. *J. Chem. Phys.,* (submitted); IBM Research Report POK-13 (August 18, 1982).

Pohl, F. J.; Baehr, W.; Holbrook, J., 1972. *Proc. Natl. Acad. Sci., USA, 69,* 3805.

Popkie, H. Kistenmacher, H.; Clementi, E., 1973. *J. Chem. Phys., 59,* 1825.

Queran, M.; Weisbuch, G., 1980. *Biopolymers, 19,* 353.

Ragazzi, R.; Ferro, D.; Clementi, E., 1979. *J. Chem. Phys., 70,* 1040.

Record, M. T., 1967. *Biopolymers, 5,* 975.

Romano, S.; Clementi, E., 1980. *Int. J. Quantum Chem., 17,* 1007.

Schellman, J.; Stigter, D., 1977. *Biopolymers, 16,* 1415.

Schildkraut, C.; Lifson, S., 1965. *Biopolymers, 3,* 195.

Scordamaglia, R.; Cavallone, F.; Clementi, E., 1977. *J. Am. Chem. Soc., 99,* 5545.

Scrocco, E.; Tomasi, J., 1978. *Top. Curr. Chem., 42,* 95.

Sundaralingam, R., 1983. In *Structure and Dynamics: Nucleic Acids and Proteins,* Clementi, E.; Sarma, R., Eds., Adenine Press, N.Y., pp. 135-148.

Texter, J., 1978. *Prog. Biophys. Molec. Biol., 33,* 83.

Wang, A. H.; Fujii, S.; van Boom, J. H.; Rich, A., 1982. *Proc Nat. Acad. Sci. USA, 79,* 3968.

Wang, A. H.; Quigley, G. J.; Kolpak, F. J.; van der Marel, G.; van Boom, J. H.; Rich, A., 1980. *Science, 211,* 171.

Watts, R. O.; Clementi, E.; Fromm, J., 1974. *J. Chem. Phys., 61,* 2550.

Welti, N.; Pretsch, E.; Clementi, E.; Simon, W., 1982. *Helv. Chim. Acta., 65,* 1996.

Zimmerman, S.; Peiffer, B., 1980. *J. Mol. Biol., 142,* 315

Lecture 5

Water Structure in the Gramicidin A Transmembrane Channel

S. L. Fornili*, D. P. Vercauteren and E. Clementi
IBM Corporation
Dept. D55/Bldg. 701-2
P.O. Box 390
Poughkeepsie, New York 12602

Abstract

The interaction energy and the structure of water molecules either inside the Gramicidin A transmembrane channel or at its two extremities is examined with the use of iso-energy maps and Monte Carlo simulations. The shape of the channel as experienced by water is analyzed in detail. Variations in the hydration structure due to the presence of a Na+ ion placed at several positions along the channel are simulated, analyzed and discussed. Preliminary data on Li+ and K+ interacting with Gramicidin A and the system of water molecules are reported. The Gramicidin A atomic coordinates have been taken from Urry's recent papers.

Introduction

The primary function of a biological membrane is the selective control on the flow of chemicals. As it is well known, simple inorganic cations, for example, Na+ and K+, can be transported across membranes; the difference in ionic concentration at the two sides of the membrane brings about electrical potential differences of basic importance especially in neuron's firing membranes. One of the possible mechanisms for an ion to cross a membrane is migration through a molecular transmembrane channel. The peptide Gramicidin A, GA, can act as a molecular channel specifically for monovalent cations. This notable discovery by Hladky and Haydon (1) was soon followed by proposals on the structure of GA; in this context, we recall works by Urry (2), Ramachandran and Chandrasekharan (3), Veatch (4), Ovchinnichov (5), and Lotz (6). Recent accounts of two somewhat different proposals for the structural representation of a GA channel can be found in publications by Ovchinnichov (7) and Urry (8).

This paper will analyze the structure of water molecules interacting with a GA channel modeled according to Urry's atomic coordinates (8). It is our pleasure to

acknowledge Professor D. A. Urry for having provided us with an updated list of the atomic coordinates of GA, prior to their publication.

We recall that the primary structure of Gramicidin A is: HCO-L•VAL[1]-GLY[2]-L•ALA[3]-D•LEU[4]-L•ALA[5]-D•VAL[6]-L•VAL[7]-D•VAL[8]-L•TRP[9]-D•LEU[10]-L•TRP[11]-D•LEU[12]-L•TRP[13]-D•LEU[14]-L•TRP[15]-$NHCH_2CH_2OH$. In Fig. 1 we report three-dimensional representations of GA. In the top-left inset, GA is projected onto the xz plane (the Z axis is the long axis of the channel); from the top-right inset (a projection into the xy plane) one can see clearly the cavity constituting the channel (the cross gives the projection of the Z axis) and the amino acid residues disposed around the two monomer backbones. The bottom inset of Fig. 1 shows the 15 residues of only one monomer; we have used different shading to allow a clear visualization. To facilitate discussions following this paper, we have explicitly labeled the 15 amino acid residues of the A monomer.

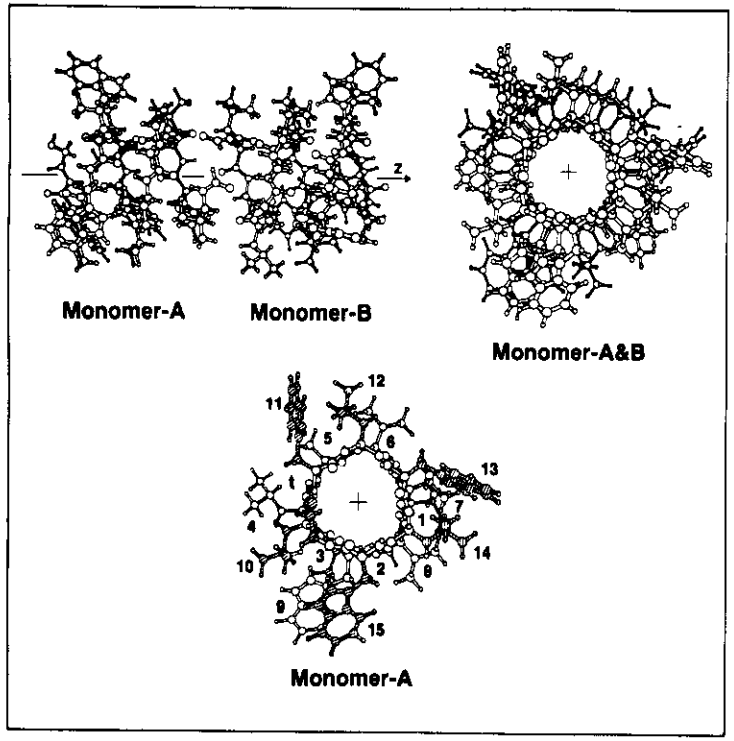


Figure 1. Top: projections of GA onto the xz plane (left) and xy plane (right); Bottom: projection of the **A** monomer of GA onto the xy plane with labels for the 15 residues.

In this work we shall discuss, 1) the number of water molecules which can be placed inside the GA channel, 2) the interaction energy of the water molecules with the channel and among themselves, and 3) the structure of the water molecules at the two extrema of the channel. In addition, we shall analyze the structural variations for the water molecules which follow the insertion of a monovalent cation at different positions along and nearby the channel. The cations we shall consider are the monovalent cations Li+, Na+, and K+. The above discussions are based on data obtained from Monte Carlo computer experiments (9,10).

Before discussing the system GA hydrated by *many* molecules of water at room temperature, we shall analyze GA interacting with *one* water molecule placed at many positions relative to GA at zero temperature. In this way we shall discuss "typical" hydration sites of GA (11); this analysis is performed by using iso-energy contour maps. As known, the iso-energy contour maps yield only a preliminary, however notably useful representation even if it cannot be related directly to any laboratory experiment. Realism in the modeling is restored with the Monte Carlo simulation where non-zero temperature, water-water interactions, and statistical weighting factors are no longer neglected.

Introductory Discussion on the Hydration Sites in GA

A water molecule at a given position and orientation relative to GA experiences an interaction energy, E, which to an approximation of the first order can be written as the sum of all the atom-atom pair-wise interactions V(i,j) between an atom i on GA and an atom j on H_2O, namely, E(k) = V(k;i,j) (k specifies the geometrical relations between GA and H_2O). For one water molecule interacting with amino acids (10) or with a poly-peptide chain (11) atom-atom potentials have been derived from *ab initio* computations. As discussed elsewhere in detail (10), the atoms of the macromolecule are grouped into "classes". We recall that a "class" characterizes the electronic environment of an atom within the molecule. This characterization is obtained by selecting as criteria 1) the atomic number, 2) the hybridization of the atom (number of bonds), 3) the atomic net charge (12), and 4) the energy difference of the atom either when in the molecule or isolated (13). An atom i, belonging to the class a and an atom j belonging to the class b interact with an energy approximated as:

$$V(i,j;a,b) = B(a,b)/R(i,j)^{12} - A(a,b)/R(i,j)^{6} + C(a,b)q(i)q(j)/R(i,j)$$

where R(i,j) is the internuclear distance for the two atoms, q(i) and q(j) their net charges and A(a,b), B(a,b), C(a,b) are numerical constants, previously derived (14,15). The value of the gross charges for the atoms of the amino acids of GA, their atomic coordinates, Molecular Orbital Valency State energies (MOVS), and class index are given in a technical report (16), available upon request. The class index of each atom allows for easy retrieval of the value for the A, B, and C fitting constants previously tabulated in Ref. 14 and 15 and now collected in Table I. Notice that the constants reported in this table for the backbone atoms represent a

Table I
Potential coefficients for the molecule-water complexes. A(O), B(O), C(O) are the coefficients of interaction with the oxygen atom of water. A(H), B(H), C(H) are the coefficients of interaction with the hydrogen atom of water.

Cl	A(O)	B(O)	C(O)	A(H)	B(H)	C(H)
6	.1021D+02	.2946D+06	.1000D+01	.3560D+01	.3892D+05	.1000D+01
7	.1107D+03	.2765D+06	.9991D+00	.1904D+02	.6190D+04	.1000D+01
9	.1129D+02	.3053D+06	.9996D+00	.8453D+02	.8706D+03	.1000D+01
13	.3444D+03	.5394D+06	.1000D+01	.6188D+00	.2029D+05	.9959D+00
14	.1262D+03	.1304D+06	.1016D+01	.9686D+01	.1302D+06	.1001D+01
15	.1618D+04	.6263D+06	.1001D+01	.4419D+03	.1271D+05	.1001D+01
18	.1021D+02	.5205D+06	.1000D+01	.3560D+01	.4045D+05	.1002D+01
33	.3481D+03	.1169D+05	.1008D+01	.1168D+02	.1242D+06	.1001D+01
34	.7065D+01	.5700D+06	.1004D+01	.2413D+03	.2396D+04	.9996D+00
35	.2910D+01	.9539D+04	.1019D+01	.1803D+02	.1953D+04	.1003D+01
36	.4071D+02	.5190D+06	.1000D+01	.8305D+02	.9119D+04	.9998D+00
37	.2721D+03	.3382D+04	.9994D+00	.4256D+01	.2040D+04	.1000D+01
38	.3349D+02	.3253D+06	.9999D+00	.1383D+03	.4837D+04	.9992D+00
39	.1132D+02	.5448D+04	.1001D+01	.1467D+01	.9434D+03	.1000D+01
40	.2442D+03	.4360D+04	.1001D+01	.1126D+01	.6979D+05	.1001D+01
41	.1817D+01	.1986D+06	.1000D+01	.1071D+03	.1617D+04	.1000D+01
Li+	−.5095D+03	.1023D+04	.1328D+01	.2271D+03	.2718D+04	.1303D+01
Na+	−.4427D+03	.2093D+05	.1129D+01	−.1198D+03	.0000D+00	.1097D+01
K+	.0000D+00	.2677D+06	.1139D+01	−.3661D+03	.3698D+04	.1125D+01

refinement of those given in Ref. 15. In the bottom three lines of the table, we report the fitting constants describing the Li+, Na+, and K+ interactions with water. It is our pleasure to acknowledge that the constants have been made available to us by Dr. G. Corongiu. We recall that the gross charges reported in the table are obtained with Mulliken's algorithm (12) from *ab initio* Self Consistent Field computations using a minimal basis set which has been reported elsewhere (17), consisting of seven s-type gaussian functions and three 2p-type gaussian functions. Earlier computations (14,15) used three s-type functions for the hydrogen atoms. In the last few years we have used four s-type functions, mainly to decrease the basis set superposition error. These computations have been performed with a computer program discussed previously (18).

The position of the hydration sites of GA can be obtained rather easily by probing with one water molecule in or nearby the channel. Because of the electrostatic charge distributions of the water and the Gramicidin A molecules, for a given position of the water molecule relative to GA, there corresponds, in general, only one optimal orentation for the water's OH bonds. Computer programs written to obtain the interaction energy for an optimally oriented water molecule with the oxygen atom placed at the grid points on a predetermined plane are available (19). These programs provide the input data needed to obtain the iso-energy maps below analyzed.

In Fig. 2, we report iso-energy maps for the L·ALA, L·VAL, D·VAL, D·LEU, L·TRP residues, the ethanolamine (HOC-NH-CH_2-CO-NH-CH_2-CH_2OH) group and for the two central formyl heads. We have selected one of the many possible planes cutting the molecules; the selected grid mesh is 0.5 a.u. Each map is reported twice. The one in two dimensions contains an ORTEP projection of the amino acid into the selected plane (the residues are shaded compared to the peptide backbone for a clearer visualization). The second representation is three-dimensional; it shows very explicitly both the attractive site minima and the repulsive regions which have the appearance of "mesa" because the repulsive energies have been cut off at 5 kcal/mol. The contour to contour interval is 1 kcal/mol. Only a few of the contours and the minima are labeled to allow easy reading of the interaction energy values. As discussed elsewhere in more detail (11) a hydration site might exist when there is a minimum (see for example in Fig. 2). This is often the case when the minimum is deeper than about 6 kcal/mol, namely, deeper than the interaction energy between two water molecules. If the GA-water interaction energy at the minimum is less than about 6 kcal/mol, then the water-water interaction might be large enough to become a dominant factor. In this case, no correspondence may be established between the existence of a minimum in the iso-energy map (see Fig. 2) and of a

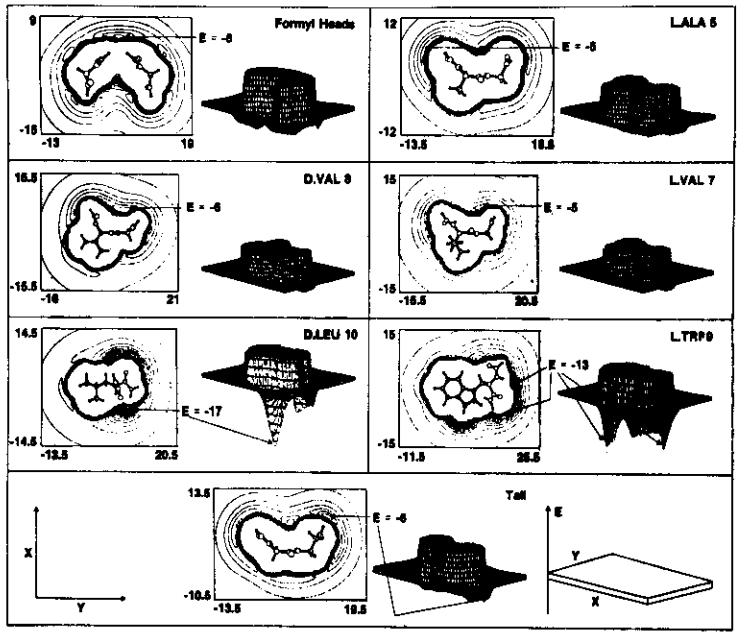


Figure 2. Iso-energy maps of the different GA's residues including the formyl heads and ethanolamine tail interacting with water. Coordinates (X,Y) and interaction energies (E) are given in a.u. and kcal/mol, respectively.

hydration site. Notice that a complete discussion of the hydration sites would require an analysis of the iso-energy maps for many planes. Such maps are also reported in the technical report (16).

In an equivalent way we have obtained the iso-energy contour maps for a plane bisecting the GA channel as shown in Fig. 3. The left inset reports a low resolution (the selected grid mesh is 1 a.u.) iso-energy map of the entire channel and of the regions at the two extrema. Notice that the length of the channel is from Z = 0 to about Z = ± 24 a.u., namely about 48 a.u.; we have reported the iso-energy contours expanding up to Z = ± 47 a.u. Two hard core areas border the channel. In the central and right-side inset we report detailed views (grid meshes = 0.5 a.u.); at the center we detail the area enclosed by the dotted line perimeter in the left-side inset. Notice the very prominent minima (E is given in kcal/mol) at the entrance of the channel and the sequence of energy minima along the channel. The three-dimensional "shape" of the channel can be appreciated by rotating the cross section plane by 90° and presenting the corresponding iso-energy contour map in the right inset of Fig. 3.

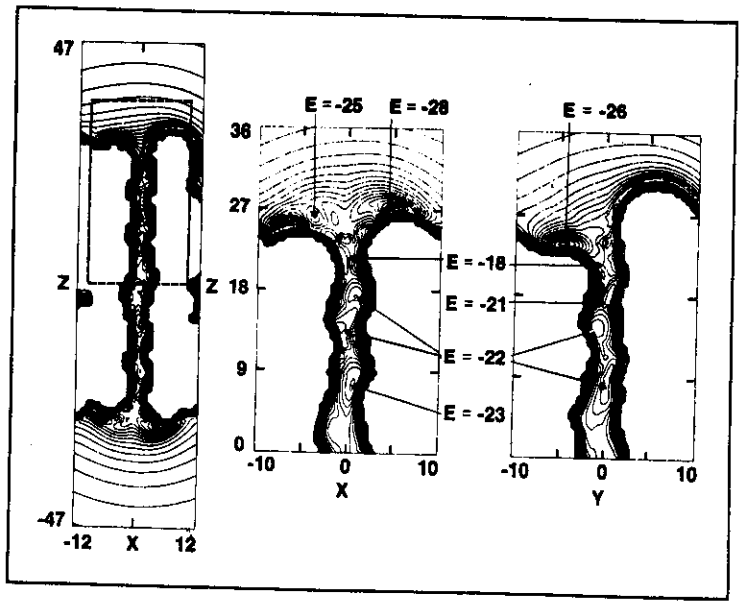


Figure 3. Left: low resolution iso-energy map for GA interacting with water (xz cross section). Center: detail of inset on the left. Right: same as center but for cross section rotated by 90° around Z axis (yz cross section). Coordinates (X,Y,Z) and interaction energies (E) are given in a.u. and kcal/mol, respectively.

It is clear and evident that the channel is very attractive to a water molecule. A water molecule, moving along the channel, will find itself in regions of variable attraction separated by energy barriers of about 3 kcal/mol. Notice, for example, the relatively less attractive region at the middle of the channel at Z = 0 a.u. These energy variations can be enhanced and time modulated if one takes into account the vibrational modes specific for GA. We refer to Urry's paper (8) for comments on the role of the librating peptide C=O groups at the channel wall. In this context we recall, that among the early computer simulations on GA, Fischer, Brickmann, and Lauger (20), did perform molecular dynamics simulations with a model where GA is approximated simply by an helical arrangement of polar carboxyl groups with properly parameterized electrostatic charges. The same model has been more recently used by Schröder, Brickmann, and Fischer (21), where qualitatively good agreement was found between molecular dynamics simulations and the diffusion constants for the ions.

A very detailed view of the channel energetic is now given in Fig. 4, where sixteen iso-energy maps are reported for xy cross sections of the channel for Z = 0, 1, ..., 15 Å (grid meshes = 0.25 a.u.). At the bottom of the figure we report a graph giving the lowest energy determined in each map. The positions a, b, c, and d in the energy profile (bottom inset), also reported in the xy cross sections (top inset), identify the energy minima; whereas a and b are outside the channel region, the minima at c and d are definitively within the channel. It is noteworthy to point out that four potential energy minima, each one separated by a barrier is postulated by Eisenman et al. (22) in order to account for the experimental transport rate (22,23). At Z = 0 Å, the attraction of GA to one molecule of water is pronounced even if it is relatively weaker than at any other Z value, an exception being made for positions very near the two ends of the channel—immediately before the most attractive sites (positions a and b)—previously noted in Fig. 3. Notice that the cross section of the channel is not circular, but assumes rather irregular shapes which gradually evolve and change with Z. Notice also that the most attractive position for water (energy minimum) is in general, not exactly at the center of the xy cross section (positions c and d in the top inset), but close to it. From these cross sections we can conclude that the channel ends at Z = 12 or Z = 13 Å; namely that the channel is 25±1 Å long. Thereafter, there is a "extuary" region which ends up at about Z = 15 or 16 Å. However, the attraction to water extends *much further*; this informs us that one cannot properly talk of "bulk water" if not for Z equal or larger than possibly ±20 or ±25 Å. In this sense the channel makes itself felt for a length of 40 to 50 Å, a value not much shorter than the standard thickness of many biological membranes.

Before reporting on the Monte Carlo simulations we shall discuss one more detail at the single water level. In Fig. 5 we consider two cylindrical iso-energy maps obtained by selecting cylinders with a radius of 1.5 and 2.0 a.u., respectively (grid meshes = 0.5 a.u.). The cylindrical maps are flattened out in order to allow for a full display. Notice that the hard core appears in the map of R = 1.5 a.u. and gets more prominent for the map of R = 2.0 a.u. As one can see by inspection of the

insets in Fig. 4, between R = 3 a.u. and R = 4 a.u. the hard core becomes the only feature. Notice in addition that the backbone of GA extends much further than R = 2.0 a.u. as evident from the right-side inset of Fig. 5. Combining the informa-

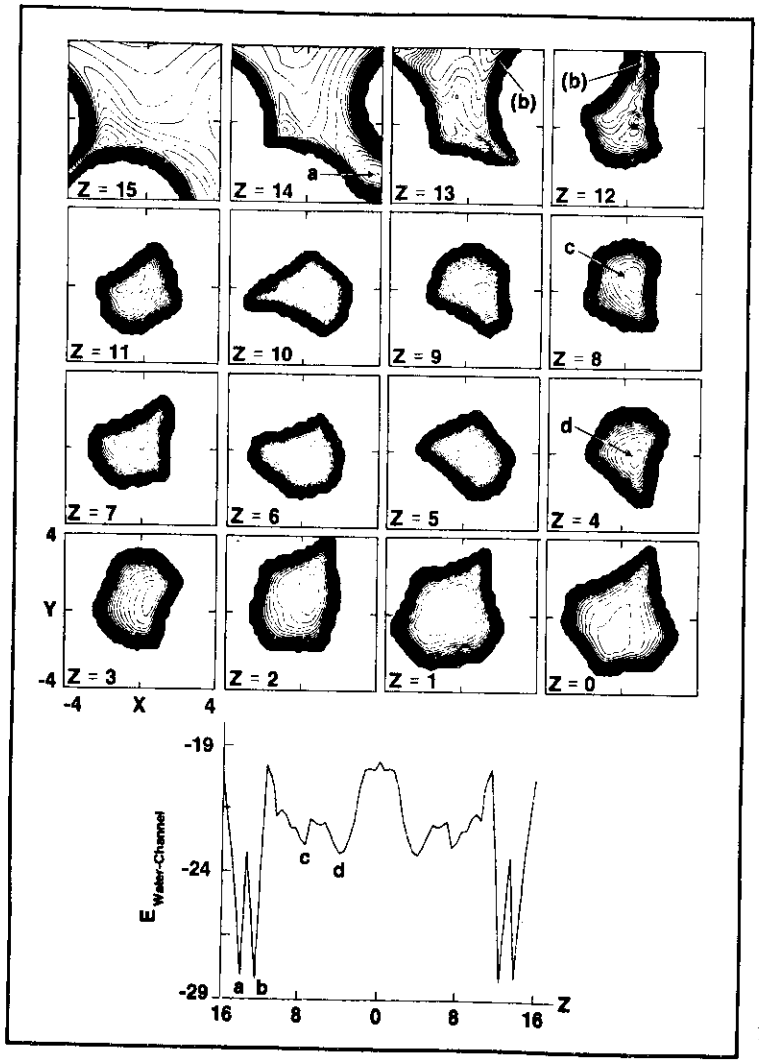


Figure 4. Sixteen cross sections of the channel from Z = 0 Å to Z = 15 Å and minimum interaction energy as function of the channel length. Coordinates (X,Y) and interaction energies (E) are given in a.u. and kcal/mol, respectively.

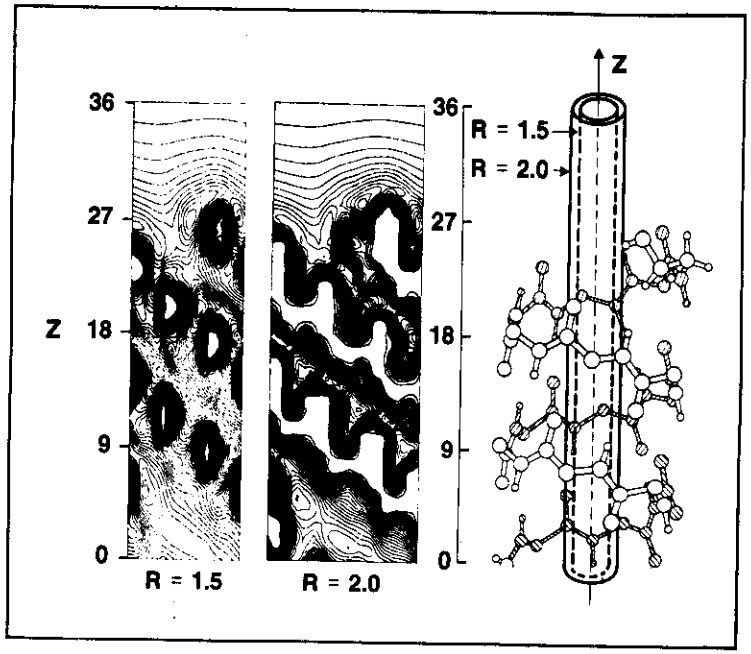


Figure 5. Cylindrical iso-energy maps flattened out: the cylinder's radii are R = 1.5 and R = 2.0 a.u., respectively. Z coordinates are given in a.u. The right-side inset is a representation of the helical backbone only.

tion of Figs. 3, 4 and 5 it seems clear that the channel is much more like a spiral than a cylinder. This is due to the helical nature of GA. Any loosening and tightening of the spiral will bring about not only a lengthening and a shortening of the channel but also the possibility of "occlusion" for the channel which is relatively very narrow with hard core protrusions extending up to R = 1.5 a.u., nearly blocking the channel. In this context it is clear that, being GA encircled by phospholipids, the dynamics of the phospholipid layers in the membrane can have profound effects on the "geometry" of the channel and hence on the ionic transport mechanism. In this paper we have neglected the interaction with the phospholipids, which will however be included in a forthcoming work (24).

Monte Carlo Simulations

In order to obtain a more realistic representation on the structure and interactions of many water molecules with GA, we must introduce 1) a temperature different from zero degree Kelvin and consequently distributions of conformations Boltzman's weighted, and 2) the interactions not only of the water molecules with GA but also

among themselves. This aim is ideally reached by performing Monte Carlo simulations as described originally in Metropolis, et al. work (9) and many times thereafter as for example in Refs. 10 and 25. Within a cylindrical volume we place a number of water molecules (about 80) surrounding GA. The volume length of the cylinder is 48 Å and the radius of the circular base is 5.5 Å; the cartesian coordinate frame is therefore the same as previously described in Fig. 3. With the above choice for the number of water molecules and the corresponding volume, the liquid water density is about one. The water molecules are constrained to stay within the cylindrical volume. The GA is constrained to remain rigid.

The Monte Carlo computer experiments we have performed are exemplified by the insets in Fig. 6, where we report one specific configuration for the water molecules positioned both inside and outside the channel (bottom inset), and the probability distribution, P(Z), for the oxygen atoms (full line) and for the hydrogen atoms (dotted line) as function of the Z axis (top inset). Notice that GA is not shown in

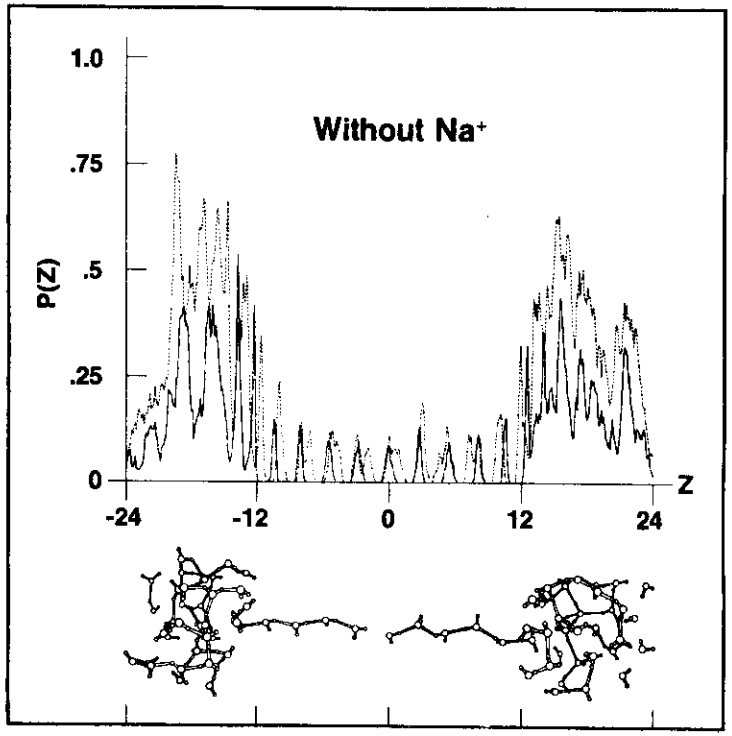


Figure 6. Top: water's oxygen (full line) and hydrogen (dotted line) probability distributions along the Z axis (in Å) of GA. Bottom: a statistically significant configuration of the water molecules.

Fig. 6 in order to simplify the figure; it occupies the volume around the long water filament, namely, the area shown as empty in the bottom inset of the figure as evident by comparing Fig. 6 with Fig. 3. In our experiments we either exclude or include a cation (Li+, Na+, or K+) which is placed at fixed positions within the cylinder. For each experiment, namely, for each position of the cation, we consider about 2×10^5 different Monte Carlo steps or configurations for the water molecules after equilibration; the latter was reached after about 4×10^5 Monte Carlo steps. We recall that each configuration is obtained from a previous one by random selection of one water molecule and by random displacement (the displacements are for distances not greater than 0.3 Å and for rotations not larger than 18 degrees). The simulated temperature is 300°K.

In the first experiment we consider a system composed of 81 water molecules and GA. As shown in the top inset of Fig. 6, the water molecules form a single file within the channel in accordance with the electrokinetic measurements of Rosenberg, et al. (26,27) and Levitt, et al. (28). *There are nine water molecules between Z = −12 Å and Z = 12 Å.* We recall that the channel is about 26 Å in length (from Z = −13 to Z = 13), the exact definition depending on where one assumes the beginning of the two outlets.

The water molecules are hydrogen bonded together as clearly shown either from the bottom inset or from the statistical distribution (top inset). We recall that we display a hydrogen bond between two water molecules whenever the O-O internuclear separation is not larger than about 3.50 Å and the O-H bond of one water molecule points toward the oxygen atom of the second molecule with an OĤO angle between 180° (linear bond) and 150° (bent hydrogen bond). An alternative way is to report all the water molecule pairs with an O-O separation smaller or equal to some threshold value corresponding either to hydrogen bonded pair or to a repulsive pair; note that the latter possibility has very low probability to occur because of Boltzman's distribution. If the threshold is very tight, then only strong hydrogen bond are selected; this is the case used in our figures. The water-water network is also strongly connected at the two channel ends. The corresponding probability distribution intensity increases immediately outside the channel, as expected. Inside the channel, due to its narrowness, one can have an accurate count of the water molecules and of the hydrogen-bonding pattern. Comparing the two insets, we notice that the water molecules of the single configuration have a structure which very closely corresponds to the probability distribution.

In the remaining Monte Carlo experiments we have placed a Na+ ion at different fixed positions along the Z axis. The positions are tabulated in Table II where we report the water-water, the GA-water and the ion-water average interaction energies in kJ/mol (all data are *per-water* molecule).

Fig. 7 corresponds to the experiment where we have fixed a Na+ at X = 0, Y = 0, and Z = 0 Å. The top inset again reports the statistical distribution of the oxygen and hydrogen atoms, whereas on the bottom inset we report a specific conformation.

Comparing the two insets, it appears that the specific conformation is statistically representative. Notice that this is due to the combined effect of a) the narrowness of the channel, b) the two deep energy minima (or the two very stable water sites) at the entrance of the channel, and c) by keeping the position of the cation fixed. Notice that the last fact is likely the least important; indeed in Fig. 6 one can notice that the single conformation representation and the statistical distribution are very close one to another. In both Figs. 6 and 7, notice that there is a strong indication that the water molecules inside the channel use one O-H to make an hydrogen bond, and the second one rotates water to water because it follows the spiraling nature of the channel. The breaks (reported in the single conformation representation as well as in the statistical probability distribution) in the water filament are due to the selection of a threshold for O-O of 3.0 Å; by increasing this value to 3.5 Å the filament has no break if not at the ion.

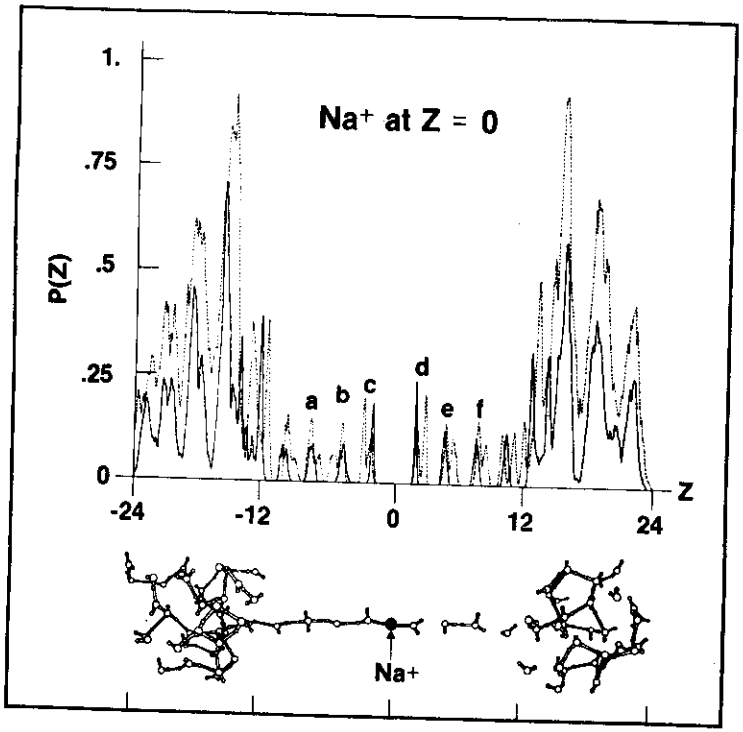


Figure 7. Top: water's oxygen (full line) and hydrogen (dotted line) probability distributions along the Z axis (in Å) of GA with a fixed Na+ ion at X = 0, Y = 0, and Z = 0 Å. Bottom: a statistically significant configuration of the water molecules (see Fig. 8 for additional details).

In Fig. 7 we report the probability distribution of the oxygen and hydrogen atoms along Z. Since there is no reason to assume that the distribution is isotropical, in Fig. 8 we report the probability distributions projected onto the xy plane for the six water molecules designated by the six peaks a, b, ..., f in Fig. 7. In order to offer as much chemical insight as possible, the probability distributions of Fig. 8 are given separately for the hydrogen atoms (at the right) and for the oxygen atoms (at the left). The insets for the oxygen atom probability display in addition information on the location where the hydrogen probability is large and on the backbone of GA with indication of the connected residue (same label as in Fig. 1). From the probabilities along Z and into the xy plane one can visualize the sterical organization of the water molecules to the left and to the right of the ion and the amount of mobility of each water molecule *even keeping the cation fixed.*

In Figs. 9 and 10 we compare four probability distributions (Fig. 9) and four corresponding statistically significant single configurations (Fig. 10) obtained when the Na+ is at Z = 14 Å, 12.5 Å, 10.63 Å and 5 Å; namely, we consider the ion immediately outside the channel end, at the very beginning, at the experimental binding site (8) close to the C=O group of L·TRP[11], and around the middle of one monomer. In presence of one fixed cation, between Z = −12 Å and Z = 12 Å, we

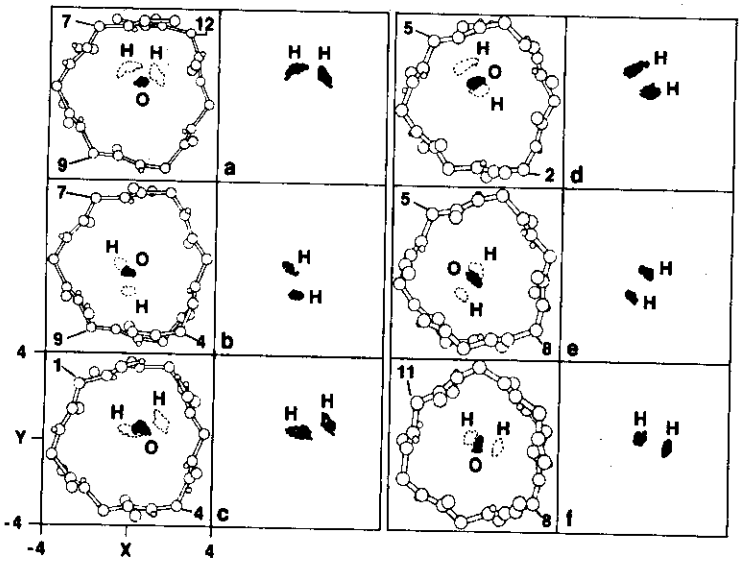


Figure 8. Oxygen and hydrogen atoms probability distribution maps projected onto the xy plane (in a.u.) for six selected regions of the GA channel corresponding to the probability peaks of the six water molecules in Fig. 7. The GA backbone atoms nearest to water molecules are shown. The number connected to the C indicates the connected residue.

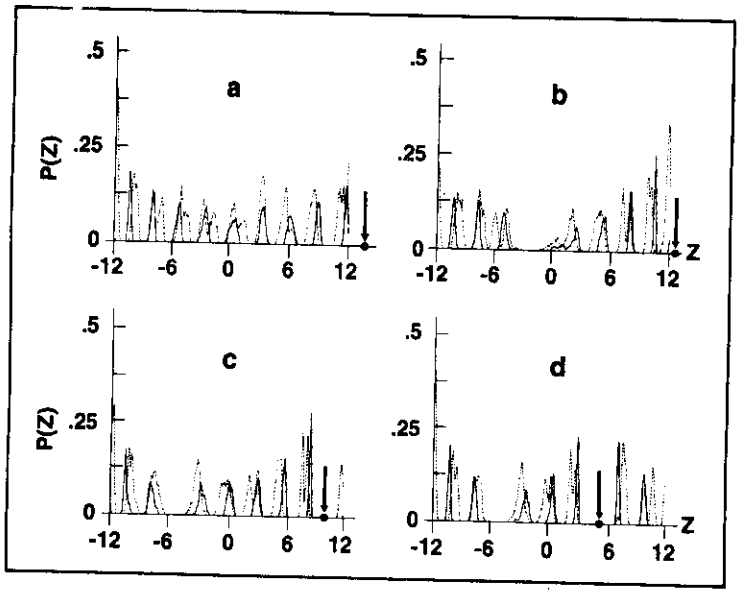


Figure 9. Water molecules probability distributions with Na+ at Z = 14(a), 12.5(b), 10.63(c), and 5 Å(d). Arrows point to the Na+ positions (see Fig. 10 for ORTEP representations).

observe 8 (Fig. 7) *or 7 water molecules* (Figs. 9 and 10) *within the channel,* an intermediate value between the number of water molecules coupled to the transport of Na+ deduced experimentally by Rosenberg (26,27) (6-7 water molecules) and the one observed by Levitt (28) (12 molecules).

Notice, also, from the Figs. 9 and 10 that when the ion is at the border of the channel it can still coordinate more than two water molecules (for Z = 14 Å and Z = 12.5 Å the Na+ coordination number of the first shell is three). When the ion reaches the binding site (10.63 Å), only two water molecules find room to solvate it. This variation in the solvation characteristic is also shown by the interaction energies of Table II, which decreases from about 10 to 5 kJ/mol. The value at Z = 24 Å should be ignored since the cation is located close to the boundary selected for our cylinder; indeed the values at Z = 24 Å are reported only to show the boundaries effects. The introduction of the cation lowers the attraction GA-water, since water is displaced in order to make room for Na+ (we could talk of "local" dehydration). The above energetic losses are compensated by the attraction GA-Na+ which is zero when Na+ is far from the channel and increases when the Na+ penetrates the channel. In a forthcoming paper we shall analyze this delicate balance in detail.

In our last experiment we compare Na+ at Z = 0 Å with Li+ or K+ at X = 0, Y = 0, and Z = 0 Å (Fig. 11). In all cases the ion, when in the channel, is

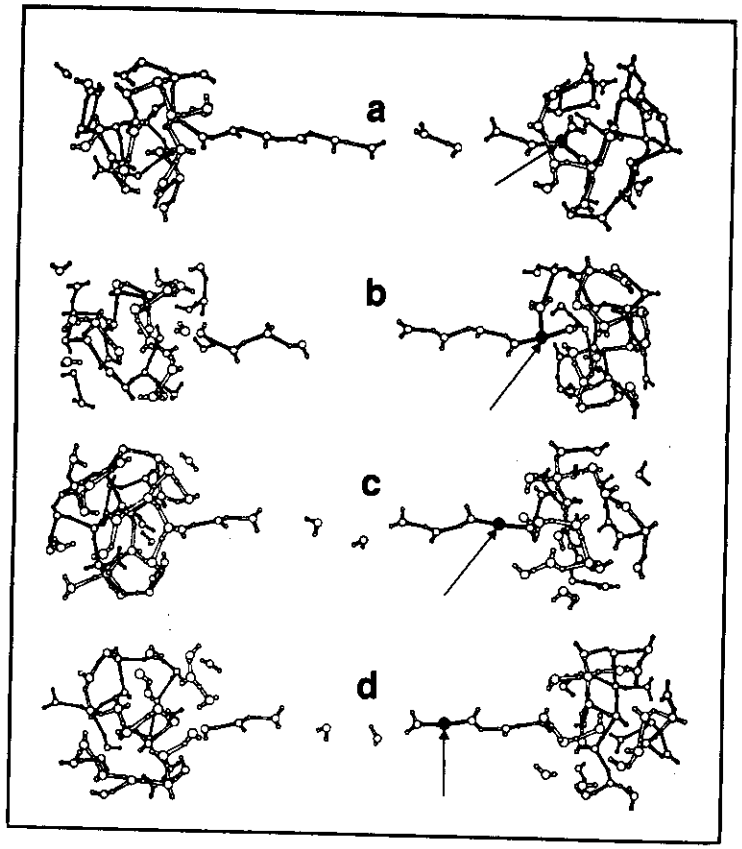


Figure 10. Statistically significant configurations for water molecules with Na+ at Z = 14(a), 12.5(b), 10.63(c), and 5 Å(d). Arrows point to the Na+ ions (see Fig. 9 for probability distributions).

coordinated to *two* water molecules; the directional interaction of the oxygen atoms of the GA carbonyl groups with the ion compensates for the loss of the coordinating water molecules of the ion in bulk water. In this sense, the coordination remains nearly constant for the ion both in the solvent water and in the GA channel. The cation-water interaction increases in the order Li+ > Na+ > K+, but these energy differences are relatively small, on the order of 1 to 2 kJ/mol, not at all as large as the differences in solvation energy in bulk water (see Table II). The water-water interaction for the experiments with Li+ and K+ are essentially the same as those found for the Na+ experiment. Therefore the main source of specificity in the GA-ion mechanism is in the differences of the interaction energy between GA and different ions (24), and also in the elastic properties of the channel

Table II
Positions (A) of the cation and computed (Monte Carlo simulations) water-water, GA-water and cation-water average interaction energies (kJ/mol; all data are per water molecule).

	E H_2O/H_2O	E GA/H_2O	E Ion/H_2O
Without ion	−25.54±0.07	−50.08±0.07	–
Na+ at Z=24.00A	−24.94±0.10	−48.23±0.04	−7.87±0.23
Na+ at Z=20.00A	−23.29±0.11	−49.42±0.05	−10.02±0.05
Na+ at Z=14.00A	−25.26±0.06	−47.59±0.08	−8.22±0.04
Na+ at Z=12.50A	−25.55±0.08	−47.69±0.05	−7.63±0.06
Na+ at Z=11.20A	−25.82±0.05	−47.76±0.06	−6.73±0.05
Na+ at Z=10.63A	−25.93±0.11	−48.40±0.05	−6.33±0.04
Na+ at Z=10.00A	−26.26±0.07	−47.79±0.05	−5.89±0.03
Na+ at Z= 7.50A	−26.30±0.11	−47.79±0.06	−5.10±0.02
Na+ at Z= 5.00A	−26.24±0.09	−47.98±0.07	−5.13±0.02
Na+ at Z= 2.50A	−25.88±0.06	−48.40±0.08	−4.86±0.02
Na+ at Z= 0.00A	−26.69±0.06	−48.35±0.07	−4.94±0.02
Li+ at Z= 0.00A	−25.49±0.12	−49.21±0.04	−6.04±0.03
K+ at Z= 0.00A	−26.01±0.09	−48.84±0.06	−3.01±0.02

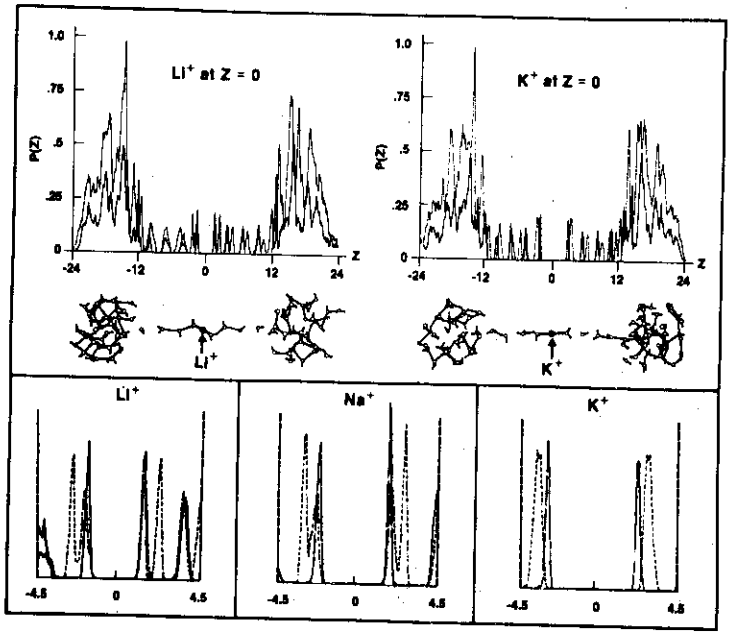


Figure 11. Top: Water's oxygen (full line) and hydrogen (dotted line) probability distributions along the Z axis (in Å) of GA with Li+ or K+ at X = 0, Y = 0, and Z = 0 Å. Bottom: detailed distributions for the two water molecules coordinating Li+, Na+, and K+ in GA at X = 0, Y = 0, and Z = 0 Å.

(see for example Refs. 8 and 21). These two factors require rather detailed knowledge of the interaction energy potentials between an ion and the *full* GA channel and of the influence of the phospholipids on the GA dynamics.

Acknowledgment

One of us (D.P.V.) would like to thank the National Foundation for Cancer Research (NFCR) for its financial support.

References and Footnotes

1. *Hladky, S. B. and Haydon, D. A., Nature, 225,* 451-453 (1970).
2. Urry, D. W., *Proc. Nat. Acad. Sci. USA, 68,* 672-676 (1971).
3. Ramachandran, G. N. and Chandrasekharan, R., *Ind. J. Biochem. Biophys., 9,* 1-11 (1972).
4. Veatch, W. R., Fossel, E. T. and Blout, E. R., *Biochemistry, 13,* 5249-5256 (1974).
5. Fossel, E. T., Veatch, W. R., Ovchinnikov, Yu. A. and Blout, E. R., *Biochemistry, 13,* 5264-5275 (1974).
6. Lotz, B., Colonna-Cesari, F., Heitz, F. and Spach, G., *J. Mol. Biol., 106,* 915-942 (1976).
7. Ovchinnikov, Yu. A. and Ivanov, V. T., in *Conformation in Biology,* (Srinivasan, R. and Sarma, R. H., eds.) Adenine Press, N.Y., (1983), pp. 155-174.
8. Urry, D. W., Trapane, T. L. and Prasad, K. U., *Int. J. Quant. Chem., Quantum Biol. Symposium, 9,* 31-40 (1982).
9. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E., *J. Chem. Phys., 21,* 1087-1092 (1953).
10. For a general discussion on hydration of biomolecules see for example, Clementi, E., in *Lecture Notes in Chemistry, Vol 19,* "Computational Aspects for Large Chemical Systems", Springer-Verlag, Heidelberg, Berlin, N.Y., (1980).
11. For a discussion on the characterization of an hydration site see for example, Clementi, E., in *Structure and Dynamics: Nucleic Acids and Proteins* (Clementi, E. and Sarma, R. H., eds.) Adenine Press, N.Y., (1983), pp. 321-364.
12. Mulliken, R. S., *J. Chem. Phys., 23,* 1833-1840 (1955).
13. Clementi, E., Corongiu, G. and Ranghino, G., *J. Chem. Phys., 74,* 578-588 (1981).
14. Clementi, E., Cavallone, F. and Scordamaglia, R., *J. Amer. Chem. Soc., 99,* 5531-5557 (1977).
15. Ragazzi, M., Ferro, D. R. and Clementi, E., *J. Chem. Phys., 70,* 1040-1050 (1979).
16. Fornili, S. L., Vercauteren, D. P. and Clementi, E., *IBM Res. Report POK-24,* June (1983).
17. Gianolio, L., Pavani, R. and Clementi, E., *Gazz. Chim. Ital., 108,* 181-205 (1978).
18. See for example, a) Davis, R. and Clementi, E., *IBM Tech. Report,* December (1965); b) Ortoleva, E., Castiglione, G. and Clementi, E., *Computer Phys. Communications, 19,* 337-357 (1980).
19. Bolis, G. and Ranghino, G., Istituto Donegani, *Tech. Rep. DDC-780* (1977).
20. Fischer, W., Brickmann, J. and Lauger, P., *Biophys. Chem., 13,* 105-116 (1981).
21. Schroder, S., Brickmann, J. and Fischer, W., *Mol. Phys., 11,* 1-11 (1983).
22. Eisenman, G., Sandblom, J. and Hagglund, J., in "Electrical behaviour of single-filing channels" in *Structure and Function of Exitable Cell,* (Adelman, W., Chang, D., Leuchtag, R. and Tasaki, I., eds), Plenum Press, N.Y., (1983).
23. Lauger, P., *Biochem. Biophys. Acta, 311,* 423 (1973).
24. Fornili, S. L., Vercauteren, D. P. and Clementi, E. (to be published).
25. a) Valleau, J. P. and Whittington, S. G., in *Statistical Mechanics. Part A: Equilibrium Techniques* (Berne, B. J., ed.), Plenum Press., N.Y., (1977), pp. 137-168; Valleau, J. P. and Torrie, G. M. Idem, pp. 169-191.
26. Rosenberg, P. A. and Finkelstein, A., *J. Gen. Physiol., 72,* 327-340 (1978).
27. Rosenberg, P. A. and Finkelstein, A., *J. Gen. Physiol., 72,* 341-350 (1978).
28. Levitt, D. G., Elias, S. R. and Hautman, J. M., *Biochem. Biophys. Acta, 512,* 436-451 (1978).

Date Received: July 10, 1983 on computer tape as EBCDIC images